



TUGAS AKHIR - RE 141581

**IMPLEMENTASI MODEL STL SEASONAL TREND
DECOMPOSITION BASED ON LOESS DAN ARIMA UNTUK
PREDIKSI KONSENTRASI KUALITAS UDARA**

**AHMAD HARITSAH
3311 100 113**

**DOSEN PEMBIMBING
Dr. Eng. Arie Dipareza Syafei, ST., MEPM**

**JURUSAN TEKNIK LINGKUNGAN
Fakultas Teknik Sipil dan Perencanaan
Institut Teknologi Sepuluh Nopember
Surabaya 2015**



FINAL PROJECT - RE 141581

**APPLICATION OF A SEASONAL TREND
DECOMPOSITION PROCEDURE BASED ON LOESS
MODEL AND ARIMA FOR PREDICTING AIR
POLLUTANT CONCENTRATION**

**AHMAD HARITSAH
3311 100 113**

**SUPERVISOR
Dr. Eng. Arie Dipareza Syafei, ST., MEPM**

**DEPARTMENT OF ENVIRONMENTAL ENGINEERING
Faculty of Civil Engineering and Planning
Institut Teknologi Sepuluh Sepuluh Nopember
Surabaya 2015**

IMPLEMENTASI MODEL STL (SEASONAL TREND DECOMPOSITION BASED ON LOESS) DAN ARIMA UNTUK PREDIKSI KONSENTRASI KUALITAS UDARA

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Teknik
pada
Program Studi S-1 Jurusan Teknik Lingkungan
Fakultas Teknik Sipil dan Perencanaan
Institut Teknologi Sepuluh Nopember

Oleh :

AHMAD HARITSAH

NRP. 3311100113

Disetujui oleh

Pembimbing Tugas Akhir:



Dr. Eng. Arie Dipareza Syafei, ST., MEPM

NIP. 19820119 200501 1 001



ABSTRAK

Implementasi Model *STL (Seasonal Trend Decomposition Procedure based on Loess)* dan *ARIMA* untuk Prediksi Konsentrasi Kualitas Udara

Nama Mahasiswa : Ahmad Haritsah
NRP : 3311100113
Jurusan : Teknik Lingkungan
Dosen Pembimbing : Dr. Eng. Arie Dipareza Syafei, ST.,
MEPM

Kegiatan transportasi dan industri memberikan kontribusi sekitar 70 % terhadap pencemaran udara di kota kota besar. Beberapa faktor yang menjadi penyebab tingginya pencemaran udara dari kendaraan bermotor adalah pesatnya jumlah kendaraan bermotor, rendahnya kualitas bahan bakar minyak (BBM) serta buruknya manajemen transportasi. Pada umumnya bahan pencemar yang terdapat di perkotaan adalah SO_2 , NO_x , O_3 , CO debu dan partikulat.

Seasonal Trend Decomposition using Loess (STL) adalah suatu bentuk algoritma yang dikembangkan untuk membantu mendekomposisikan sebuah time series menjadi 3 komponen yaitu trend, musiman dan remainder (sis). Definisi dari Loess itu sendiri adalah sebuah metode regresi non parametrik. Dimana regresi tersebut memiliki keunggulan, yaitu fleksibilitas yang tinggi karena data akan dengan sendirinya membentuk estimasi kurva yang tidak dipengaruhi oleh faktor subyektif (*Outlier*)

Pada penelitian ini digunakan 2 variasi data yaitu , performa STL dan ARIMA terhadap pengaruh banyaknya data yang hilang dan performa STL dan ARIMA terhadap perbedaan data *in sample* yang *overlapping* . Model STL merupakan model yang lebih modern dibandingkan dengan ARIMA, tetapi berdasarkan hasil penelitian didapat kesimpulan bahwa model STL hanya. sesuai untuk memprediksi CO dan SO_2 . dan performa kedua model dipengaruhi oleh banyaknya data yang hilang.

Kata Kunci : ARIMA, Konsentrasi Polutan, STL

ABSTRACT

Application of a Seasonal Trend Decomposition Procedure Based on Loess Model and ARIMA for Predicting Air Pollutant Concentration

Student ID : Ahmad Haritsah
NRP : 3311100113
Departement : Environmental Engineering
Supervisor : Dr. Eng. Arie Dipareza Syafei, ST.,
MEPM

Transportation activities and industry contributed about 70 % to air pollution in big cities .Several factors that into a reason for high air pollution from motor vehicles is rapid expansion of the number of vehicle , the low quality of fuel oil (bbm) bad management and transportation .In general material pollution which there are in urban areas is SO₂, of NO_x , O₃ , CO and particulate dust (PM₁₀)

Seasonal trend of decomposition using loess (STL) is an algorithm which decomposition time series into three components which is the trend, seasonal and remainder .The definition of loess is a method of regression .The algorithm has flexibility regression and simple design. Robust Trend and seasonal components that are not distorted by transient aberrant behavior in the data.

On this research used two variations. There are Performance against the influence of the many losing data and performance against differences in sample (overlapping) .STL model is a model that more modern compared with arima, but based on the research Model STL can predict CO and SO₂. The second variation prove that losing data can influence performance model

Keyword : ARIMA, Air Pollution, STL

KATA PENGANTAR

Puji syukur penulis ucapkan ke hadirat Allah SWT, atas limpahan rahmat dan hidayahNya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul “ **Implementasi Model STL (Seasonal Trend Decomposition Procedure based on Loess) dan ARIMA untuk Prediksi Konsentrasi Kualitas Udara**”

Penulis tidak lupa mengucapkan terima kasih yang sebesar besarnya kepada :

1. Bapak Dr. Eng. Arie Dipareza Syafei, ST., MEPM, selaku dosen pembimbing yang telah sabar membimbing dan memberikan banyak masukan kepada penulis.
2. Bapak Ir. Rachmat Boedisantoso, MT. dan Bapak Abdu Fadli Assomadi, SSi., MT. selaku dosen penguji yang telah memberikan banyak masukan dalam penyusunan laporan tugas akhir.
3. Ir. Eddy Setiadi Soedjono, Dipl., SE., MSc., Ph.D, selaku Kepala Jurusan Teknik Lingkungan FTSP-ITS.
4. Dr. Ellina S Pandebesie, ST., MT., Selaku Ketua Program studi SI Teknik Lingkungan FTSP- ITS.
5. Bapak Mohammad Toha Erpan dan Ibu Gati Mareta yang selalu memberikan dukungan yang tiada henti.
6. Amru Wibowo Rahman dan Amelia Rahmatika sebagai kakak penulis yang membantu dalam memberikan saran-saran positif dan dukungan doa.
7. Izhar Rahman Dwiputra, Farid Andhika, Rizki Mulyansyah, Ghifari, Raditya Arya, Bagus Wicaksana, Devi Apriyanti dan Nurul Ramadhan sebagai tim lab udara yang bersama-sama menyelesaikan tugas akhir ini.
8. Teman-teman angkatan 2011 yang telah banyak membantu, menghibur, memberi dukungan moral serta doa, dan mau mendengarkan keluh kesah penulis.

9. *Seluruh dosen dan karyawan jurusan Teknik Lingkungan FTSP-ITS, atas ilmu dan bantuan yang telah diberikan selama penulis berada di bangku kuliah S1.*

10. *Semua pihak yang tidak dapat disebutkan satu-persatu yang telah membantu hingga pelaksanaan Tugas Akhir ini dapat terselesaikan dengan baik.*

Penulis menyadari bahwa dalam penyusunan laporan tugas akhir ini jauh dari sempurna. Penulis berharap laporan tugas akhir ini bermanfaat dan dapat menjadi dasar untuk pengembangan ilmu pengendalian pencemaran udara di masa mendatang. Penulis berharap semoga Allah SWT. memberikan imbalan yang setimpal pada mereka yang telah memberikan bantuan dan dapat menjadikan semua bantuan ini sebagai amal ibadah, Amiin Yaa Robbal 'Alamiin.

Surabaya, Mei 2015

Penulis

DAFTAR ISI

| | |
|---|----|
| ABSTRAK..... | I |
| ABSTRACT | I |
| KATA PENGANTAR..... | I |
| DAFTAR TABEL..... | V |
| DAFTAR GAMBAR | IX |
| BAB 1 PENDAHULUAN..... | 1 |
| 1.1. Latar Belakang | 1 |
| 1.2. Rumusan Masalah..... | 2 |
| 1.3. Tujuan..... | 3 |
| 1.4. Manfaat..... | 3 |
| 1.6 Ruang Lingkup | 3 |
| BAB 2 TINJAUAN PUSTAKA..... | 5 |
| 2.1. Pencemaran Udara | 5 |
| 2.1.1. Pengertian Pencemaran Udara..... | 5 |
| 2.1.2. Monitoring Zat Pencemar Udara | 6 |
| 2.2. Gas Pencemar Udara..... | 8 |
| 2.2.1. Karbon Dioksida | 8 |
| 2.2.2. Sulfur Dioksida | 9 |
| 2.2.3. Nitrogen Dioksida | 11 |
| 2.2.4. Ozon | 12 |
| 2.3. Partikulat Pencemar Udara | 14 |
| 2.4. Peramalan Deret Waktu (Time Series Forecasting)..... | 14 |
| 2.5. Software Statistik R | 22 |
| 2.5.1. Paket Library | 23 |

| | |
|---|------------|
| BAB 3 METODE PENELITIAN | 25 |
| 3.1. Gambaran Umum | 25 |
| 3.2. Lokasi dan Lingkungan Sekitar SUF | 25 |
| 3.3. Kerangka Penelitian | 26 |
| 3.4. Studi Literatur | 28 |
| 3.5. Penentuan Skenario Model | 28 |
| 3.6. Pengumpulan Data | 30 |
| 3.7. Simulasi Model | 32 |
| 3.8. Hasil dan Pembahasan | 36 |
| 3.9. Kesimpulan dan Saran | 36 |
| BAB 4 ANALISA DAN PEMBAHASAN | 37 |
| 4.1. Performa Model STL dan Model ARIMA terhadap Perbedaan Data In Sample | 37 |
| 4.1.1. SO ₂ | 37 |
| 4.1.2. CO | 58 |
| 4.1.3. PM ₁₀ | 79 |
| 4.1.4. O ₃ | 100 |
| 4.1.5. NO ₂ | 121 |
| 4.2. Performa Model STL dan Model ARIMA terhadap data yang hilang | 142 |
| 4.2.1. SO ₂ | 142 |
| 4.2.2. CO | 144 |
| 4.2.3. PM ₁₀ | 146 |
| 4.2.4. O ₃ | 148 |
| 4.2.5. NO ₂ | 150 |
| BAB 5 KESIMPULAN DAN SARAN | 153 |
| DAFTAR PUSTAKA | 155 |

DAFTAR GAMBAR

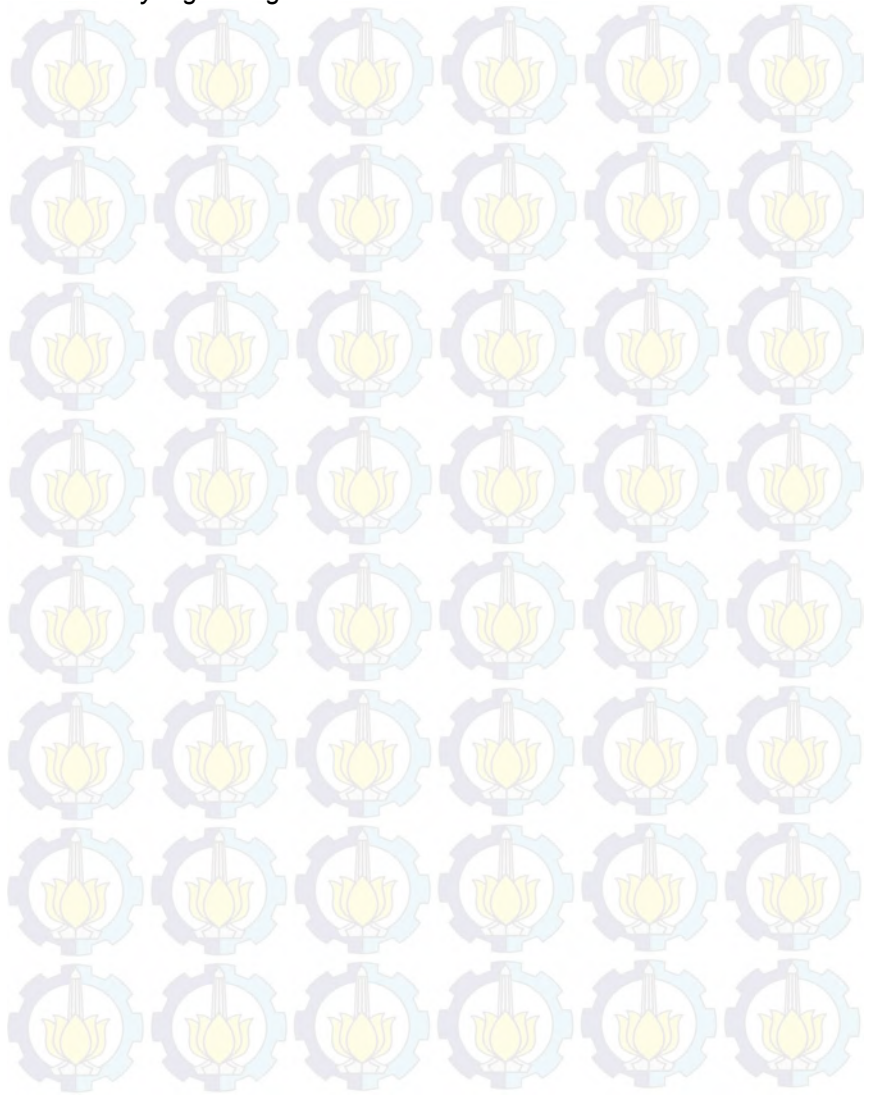
| | |
|--|-----|
| Gambar 4.1 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter SO_2 | 43 |
| Gambar 4.2 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter SO_2 | 51 |
| Gambar 4.3 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter SO_2 | 57 |
| Gambar 4.4 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter CO | 64 |
| Gambar 4.5 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter CO | 71 |
| Gambar 4.6 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter CO | 78 |
| Gambar 4.7 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter PM_{10} | 85 |
| Gambar 4.8 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter PM_{10} | 92 |
| Gambar 4.9 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter PM_{10} | 99 |
| Gambar 4.10 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter O_3 | 106 |
| Gambar 4.11 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter O_3 | 113 |
| Gambar 4.12 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter O_3 | 121 |
| Gambar 4.13 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter NO_2 | 128 |
| Gambar 4.14 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter NO_2 | 134 |
| Gambar 4.15 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter NO_2 | 141 |

DAFTAR TABEL

| | |
|---|------------|
| <i>Tabel 2. 1 Pengaruh Gas SO₂ Terhadap Manusia.....</i> | <i>10</i> |
| <i>Tabel 2.2 Transformasi Box-Cox</i> | <i>20</i> |
| <i>Tabel 4.1 Performa Model (STL)</i> | <i>37</i> |
| <i>Tabel 4.2 Performa Model (ARIMA).....</i> | <i>40</i> |
| <i>Tabel 4.3 Performa Model ARIMA dan STL</i> | <i>42</i> |
| <i>Tabel 4.4 Performa Model (STL).....</i> | <i>44</i> |
| <i>Tabel 4.5 Performa Model (ARIMA).....</i> | <i>47</i> |
| <i>Tabel 4.6 Performa Model ARIMA dan STL</i> | <i>50</i> |
| <i>Tabel 4.7 Performa Model (STL).....</i> | <i>52</i> |
| <i>Tabel 4.8 Performa Model (ARIMA).....</i> | <i>54</i> |
| <i>Tabel 4.9 Performa Model ARIMA dan STL</i> | <i>56</i> |
| <i>Tabel 4.10 Performa Model (STL).....</i> | <i>58</i> |
| <i>Tabel 4.11 Performa Model (ARIMA).....</i> | <i>61</i> |
| <i>Tabel 4.12 Performa Model ARIMA dan STL</i> | <i>63</i> |
| <i>Tabel 4.13 Performa Model (STL).....</i> | <i>65</i> |
| <i>Tabel 4.14 Performa Model (ARIMA).....</i> | <i>68</i> |
| <i>Tabel 4.15 Performa Model ARIMA dan STL</i> | <i>70</i> |
| <i>Tabel 4.16 Performa Model (STL).....</i> | <i>72</i> |
| <i>Tabel 4.17 Performa Model (ARIMA).....</i> | <i>75</i> |
| <i>Tabel 4.18 Performa Model ARIMA dan STL</i> | <i>77</i> |
| <i>Tabel 4.19 Performa Model (STL).....</i> | <i>79</i> |
| <i>Tabel 4.20 Performa Model (ARIMA).....</i> | <i>82</i> |
| <i>Tabel 4.21 Performa Model ARIMA dan STL</i> | <i>84</i> |
| <i>Tabel 4.22 Performa Model (STL).....</i> | <i>86</i> |
| <i>Tabel 4.23 Performa Model (ARIMA).....</i> | <i>89</i> |
| <i>Tabel 4.24 Performa Model ARIMA dan STL</i> | <i>91</i> |
| <i>Tabel 4.25 Performa Model (STL).....</i> | <i>93</i> |
| <i>Tabel 4.26 Performa Model (ARIMA).....</i> | <i>96</i> |
| <i>Tabel 4.27 Performa Model ARIMA dan STL</i> | <i>98</i> |
| <i>Tabel 4.28 Performa Model (STL).....</i> | <i>100</i> |

| | |
|---|-----|
| <i>Tabel 4.29 Performa Model (ARIMA)</i> | 103 |
| <i>Tabel 4.30 Performa Model ARIMA dan STL</i> | 105 |
| <i>Tabel 4.31 Performa Model (STL)</i> | 107 |
| <i>Tabel 4.32 Performa Model (ARIMA)</i> | 110 |
| <i>Tabel 4.33 Performa Model ARIMA dan STL</i> | 112 |
| <i>Tabel 4.34 Performa Model (STL)</i> | 114 |
| <i>Tabel 4.35 Performa Model (ARIMA)</i> | 117 |
| <i>Tabel 4.36 Performa Model ARIMA dan STL</i> | 119 |
| <i>Tabel 4.37 Performa Model (STL)</i> | 122 |
| <i>Tabel 4.38 Performa Model (ARIMA)</i> | 124 |
| <i>Tabel 4.39 Performa Model ARIMA dan STL</i> | 127 |
| <i>Tabel 4.40 Performa Model (STL)</i> | 128 |
| <i>Tabel 4.41 Performa Model (ARIMA)</i> | 131 |
| <i>Tabel 4.42 Performa Model ARIMA dan STL</i> | 133 |
| <i>Tabel 4.43 Performa Model (STL)</i> | 135 |
| <i>Tabel 4.44 Performa Model (ARIMA)</i> | 138 |
| <i>Tabel 4.45 Performa Model ARIMA dan STL</i> | 140 |
| <i>Tabel 4.46 Performa Model STL dan ARIMA dengan Data Hilang secara Beruntun</i> | 142 |
| <i>Tabel 4. 47 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak</i> | 143 |
| <i>Tabel 4.48 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun</i> | 144 |
| <i>Tabel 4.40 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak</i> | 145 |
| <i>Tabel 4.50 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun</i> | 146 |
| <i>Tabel 4.51 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak</i> | 147 |
| <i>Tabel 4.52 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun</i> | 148 |
| <i>Tabel 4. 53 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak</i> | 149 |
| <i>Tabel 4. 54 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun</i> | 150 |

Tabel 4. 55 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak 151



BAB 1 PENDAHULUAN

1.1. Latar Belakang

Teknologi sangat dibutuhkan manusia sebagai upaya peningkatan kualitas dan kenyamanan hidup, namun sayangnya kemajuan teknologi juga menimbulkan dampak yang merugikan manusia. Salah satunya adalah pencemaran lingkungan. Pencemaran lingkungan yang terjadi salah satunya adalah pencemaran udara. Sumber pencemaran ini berasal dari kendaraan bermotor, industri serta rumah tangga. Peningkatan sumber pencemar mengakibatkan turunnya kualitas udara ambien, oleh karena itu dibutuhkan upaya pengendalian pencemaran udara berupa pemantauan kualitas udara.

Salah satu usaha yang dapat dilakukan untuk mengatasi masalah pencemaran udara adalah dengan dilakukannya prediksi kualitas udara. Prediksi ini bertujuan untuk memprediksi konsentrasi pencemar agar masyarakat dapat melakukan tindakan pencegahan. Prediksi kualitas udara dalam jangka pendek menggunakan teknik stasistik yang tidak mempertimbangkan proses fisik dan kimia. Melainkan menggunakan data dari masa lalu (Kandya, A., dan Mohan, M., 2009). Menurut Lee, M. H dkk., 2009) metode yang baik digunakan untuk memantau dan mengendalikan kondisi kualitas udara adalah metode Time Series (deret waktu)

Analisa deret waktu adalah suatu analisa dengan menggunakan teknik statistik melalui pengoperasian model yang menggunakan data dari masa lalu untuk memprediksi masa depan (Wei, W. W.S., 2006) Data-data yang dikumpulkan pada analisis deret waktu harus berdasarkan urutan waktu, bisa dalam jam, hari, minggu, bulan, kuartal dan tahun (Damayanti, 2008). Pengukuran data deret waktu dapat dilakukan terus menerus melalui urutan waktu atau pengambilan satu set waktu secara terpisah (Suhartono, 2008)

Metode Yang sering digunakan dalam analisa deret waktu adalah AR (Autoregresif) atau metode gabungan dari AR yakni ARIMA (Autoregresif Integrated Moving Average). Model – model

tersebut secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari deret waktu (time series) secara statistik berhubungan satu sama lain (dependent). (Hendranata, 2003)

Dalam Kasus ini digunakan metode STL (Seasonal Trend Decomposition Procedure Based on Loess), STL merupakan Metode untuk mendekomposisi suatu data time series menjadi 3 Komponen yaitu komponen data kecenderungan (trend), komponen musiman (Seasonal), dan komponen residual (reminder). Metode Jenis ini mampu menangani semua jenis musiman, tidak hanya bulanan atau triwulan. Adapun komponen musiman (seasonal) pada metode ini diperbolehkan berubah seiring dengan waktu, serta tingkat perubahan dapat dikontrol oleh pengguna.

Penelitian Ini bertujuan untuk memprediksi konsentrasi zat pencemar dari hasil pembacaan Stasiun Pemantauan Kualitas Udara Ambien Otomatis (AQMS) di daerah Kebonsari, Surabaya. Terdapat Lima parameter yang direkam oleh Surabaya Fixed Station (SUF) di Kebonsari, Surabaya yaitu Sulfur dioksida (SO_2), Karbon Monoksida (CO), Nitrogen Dioksida (NO_2), Ozon (O_3) dan Partikulat (PM_{10}).

Berdasarkan Latar belakang tersebut, maka tugas akhir ini membahas tentang penerapan Model STL dan ARIMA dalam memprediksi konsentrasi kualitas udara dengan menggunakan data historis SUF di daerah Kebonsari, Surabaya

1.2. Rumusan Masalah

Rumusan masalah yang didapat dari latar belakang adalah:

Rumusan masalah pada penelitian ini adalah :

1. Bagaimana performa kedua model terhadap konsentrasi kualitas udara pada tiap parameter ?
2. Bagaimana pengaruh data in sampel yang hilang terhadap performa kedua Model (ARIMA dan STL) ?

1.3. Tujuan

Tujuan dari penelitian ini adalah:

1. *Menentukan model terbaik untuk memprediksi konsentrasi kualitas udara pada tiap parameter*
2. *Menentukan besar pengaruh performa model time series forecasting dengan adanya kehilangan data in sampel*

1.4. Manfaat

Penelitian ini bermanfaat untuk :

1. *Memberikan Informasi kepada masyarakat tentang bahaya yang ditimbulkan bahan pencemar udara, agar masyarakat dapat melakukan tindakan pencegahan.*
2. *Menjadi landasan kebijakan pemerintah Kota Surabaya dalam pengelolaan kualitas udara*

1.6 Ruang Lingkup

Ruang lingkup dalam penelitian tugas akhir ini mengambil daerah Kebonsari Surabaya. Selain itu, ruang lingkup dalam tugas akhir ini adalah :

- 1) *Penelitian dilakukan di Laboratorium Pencemaran Udara dan Perubahan Iklim Jurusan Teknik Lingkungan FTSP-ITS.*
- 2) *Penelitian dilakukan selama bulan Maret – Mei 2015.*
- 3) *Data yang digunakan dalam penelitian ini adalah data harian sekunder dari Indeks Sumber Pencemar Udara (ISPU) atau BLH Kota Surabaya.*
- 4) *Bahan polutan yang diprediksi adalah gas karbon monoksida (CO), nitrogen dioksida (NO₂), sulfur dioksida (SO₂), ozon/oksidan (O₃), dan partikulat PM₁₀.*
- 5) *Metode yang digunakan dalam penelitian ini adalah ARIMA (Autoregressive Integrated Moving Average) Model, dan Seasonal Decomposition Based on Loess*

6) Penelitian menggunakan platform opensource bernama Tinn-R sebagai script reader program dan R sebagai statistical computing program.dan Minitab 16.0 ssebagai software pembantu.

7) Penelitian tidak dipengaruhi oleh suhu, topografi dan kejadian-kejadian yang bersifat kebetulan (accidental).

BAB 2

TINJAUAN PUSTAKA

2.1. Pencemaran Udara

2.1.1. Pengertian Pencemaran Udara

Pencemaran udara adalah kehadiran materi yang tidak diinginkan di udara dalam jumlah cukup besar yang dapat memberikan efek berbahaya (Nevers, 2000). Materi yang tidak diinginkan tersebut memberikan dampak terhadap makhluk hidup seperti membahayakan kesehatan dan menyebabkan kerusakan lingkungan. Berdasarkan Keputusan Menteri Negara Kependudukan dan Lingkungan Hidup (KEPMEN KLH) No. Kep.02/Men-KLH/1988, definisi pencemaran udara adalah masuk atau dimasukkannya makhluk hidup, zat, energi dan atau komponen lain ke udara dan atau berubahnya tatanan udara oleh kegiatan manusia atau proses alam sehingga kualitas udara turun hingga ke tingkat tertentu yang menyebabkan udara menjadi kurang atau tidak dapat berfungsi lagi sesuai dengan peruntukannya. Pada dasarnya Udara di alam tidak pernah ditemukan bersih tanpa polutan sama sekali. Beberapa gas seperti Sulfur Dioksida (SO_2), Nitrogen Dioksida (NO_2), Karbon Monoksida (CO) dan Ozon (O_3) selalu dibebaskan ke udara sebagai produk sampingan dari proses-proses alami seperti aktivitas vulkanik, kebakaran hutan dan lain sebagainya. (Boedisantoso, 2002) membagi sumber pencemar udara Berdasarkan kedudukan sumbernya, yaitu :

- 1. Sumber bergerak (mobile source), contohnya : kendaraan bermotor, pesawat udara, kereta api, dan sebagainya.*
- 2. Sumber tidak bergerak (stationary source), contohnya : perumahan, daerah perdagangan, daerah industri, dan sebagainya.*

National Ambient Air Quality Standards (NAAQSs) telah menetapkan enam kriteria polutan udara, yaitu lima polutan primer dan satu polutan sekunder. Lima polutan primer tersebut terdiri atas partikulat dengan diameter kurang dari $10\ \mu\text{m}$ (PM_{10}), sulfur dioksida (SO_2), nitrogen dioksida (NO_2), karbon monoksida (CO),

serta timbal (Pb), sedangkan polutan sekunder yaitu ozon (O_3) (Cooper, 2002).

2.1.2. Monitoring Zat Pencemar Udara

Menurut Peraturan pemerintah Republik Indonesia No. 41 tahun 1999 tentang Pengendalian Pencemaran Udara, pengertian Udara Ambien adalah udara bebas di permukaan bumi pada lapisan troposfir yang berada di dalam wilayah yurisdiksi Republik Indonesia yang dibutuhkan dan mempengaruhi kesehatan manusia, makhluk hidup dan unsur lingkungan hidup lainnya. Pemantauan Kualitas Udara Ambien telah dilakukan oleh pemerintah Kota Surabaya dengan memasang Jaringan pemantauan.kualitas udara ambien, pemasangan tersebut merupakan perwujudan dari kesepakatan antara menteri menteri Lingkungan Hidup se – Asia Tenggara. Jumlah jaringan pemantauan tergantung dari jumlah penduduk, tingkat pencemaran dan keragaman dari zat pencemar itu sendiri. Sedangkan yang mendasari pemasangan jaringan pemantauan adalah tingkat konsentrasi pencemar, pola penyebaran pencemar dan inventarisasi emisi. Berdasarkan survey lokasi bersama Tim BAPEDAL Pusat, Tim Pemerintah Austria, Tim Pemerintah Kota Surabaya, Tim BAPEDAL Propinsi Jawa Timur pada tanggal 10-13 Maret 1999, ditetapkan lokasi penempatan Stasiun pemantauan kualitas udara ambien

Selain Stasiun pemantauan kualitas udara ambien, terdapat juga data display. Data display merupakan sebuah papan atau display yang berguna untuk menampilkan data kualitas udara hasil dari pemrosesan data. Data yang ditampilkan adalah data tertinggi dari ketiga stasiun pemantau, data ini mewakili kualitas udara Kota Surabaya

Lokasi penempatan Public Data Display adalah :

1. Depan Monumen Kapal Selam, Jl. Gubeng Pojok (Surabaya Pusat).
2. Depan BAPPEDA Propinsi Jatim, Jl. Pahlawan (Surabaya Utara).
3. Ring Road Jl. Mayjend Sungkono (Surabaya Barat).
4. Perempatan Jl. Dharmawangsa-Jl. Kertajaya (Surabaya Timur).
5. Depan BNI Graha Pangeran, Jl. A. Yani (Surabaya Selatan)

Parameter yang diukur dalam stasiun pemantau kualitas udara ambien di kota surabaya ada 16 parameter, yang terdiri dari :

1. Lima Parameter Kunci : PM₁₀, SO₂, O₃, NO₂, CO
2. Sebelas Parameter Pendukung dan meteorologi : NO, NO_x, Kecepatan Angin, (FF), Kecepatan hembusan angin (FF Boe), arah angin (DD), arah hembusan angin (DD Boe), kelembaban udara ambien, kelembaban udara container, suhu udara ambien, suhu container dan global radiasi

Sesuai dengan amanat Undang – Undang Lingkungan Hidup Nomor 23 Tahun 1997 bahwa masyarakat berhak untuk mendapatkan informasi mengenai kualitas lingkungan termasuk kualitas udara di kota Surabaya ini, maka pelaporan hasil pemantauan ini dikemas dalam bahasa yang mudah dipahami oleh masyarakat umum.

Informasi yang disebarkan berdasarkan Keputusan Menteri Negara Lingkungan Hidup Nomor : Kep-45/MENLH/10/1997 tentang Indeks Standar Pencemar Udara (ISPU). ISPU (Indeks Standar Pencemar Udara) adalah angka yang tidak mempunyai satuan yang menggambarkan kondisi kualitas udara ambien di

lokasi dan waktu tertentu, yang didasarkan pada dampak terhadap kesehatan manusia, nilai estetika dan makhluk hidup lainnya. Penggunaan ISPU sangat memudahkan masyarakat untuk mengetahui kondisi kualitas udara pada waktu tertentu karena sistem ini sangat informatif dan mudah dipahami oleh masyarakat luas.

2.2. Gas Pencemar Udara

2.2.1. Karbon Dioksida

Karbon monoksida adalah pencemar primer berbentuk gas yang tidak berwarna, tidak memiliki rasa, tidak berbau dan memiliki berat jenis yang lebih kecil dari udara serta sangat stabil dan inert di udara, mempunyai waktu tinggal 2 – 4 bulan (Purnomohadi, 1995). Karbon monoksida dalam jumlah yang berlebihan bersama beberapa gas lainnya seperti metana, akan menjadi gas rumah kaca yang dapat meningkatkan suhu permukaan bumi, karena menyebabkan radiasi gelombang panjang terperangkap (Soedomo, 2001). Sumber utama CO berasal dari kendaraan bermotor akibat pembakaran yang tidak sempurna, dan proses industri menduduki tempat kedua, sedangkan pembakaran sampah pertanian dan kebakaran hutan menduduki tempat ketiga dan keempat (Tjasjono, 1999). Semua aktivitas yang melibatkan pembakaran bahan-bahan organik merupakan sumber karbon monoksida. CO terbentuk juga dalam proses ledakan dan secara alami (Soedomo, 2001). Karbon monoksida mencapai 0,1%, maka kapasitas darah dalam pengangkutan oksigen berkurang 50 % . hal ini menyebabkan pemberian oksigen ke dalam tubuh berkurang serta berakibat berkurangnya pengelihan dan reksi fisik. Konsentrasi CO di udara mencapai 0,5 % (sekitar 8–14 ppm) menyebabkan pingsan dan kemudian dapat mengakibatkan kematian. Penyebaran CO biasanya terjadi pada lapisan pencampur yang paling bawah, dengan ruang gerak konveksi vertikal yang agak terbatas, akibat sifat CO sendiri yang berberat jenis besar bernilai 1250 gram/m³ (Soenarmo, 1996).

Keberadaan gas CO akan sangat berbahaya jika terhirup oleh manusia karena gas itu akan menggantikan posisi oksigen yang berkaitan dengan haemoglobin dalam darah. Gas Karbon

Monoksida berbahaya karena mudah berikatan dengan hemoglobin dalam darah membentuk karboksihemoglobin (COHb) yang jauh lebih kuat 200 kali dibandingkan dengan ikatan antara oksigen (Lindell, 2009). Menurut Soedomo (2001), Pada Paparan Karbon Monoksida selama delapan jam dengan konsentrasi sebesar 12 - 17 mg/m³ atau sekitar 10 – 15 ppm akan menimbulkan dampak buuruk bagi kesehatan manusia, terutama pada penderita penyakit jantung.

Pada daerah yang macet tingkat bahayanya cukup tinggi terhadap kasus keracunan. Asap rokok juga mengandung gas karbon dioksida, pada orang dewasa yang tidak merokok biasanya terbentuk karboksi haemoglobin tidak lebih dari 1 % tetapi pada perokok yang berat biasanya lebih tinggi yaitu 5 – 10 %. Pada wanita hamil yang merokok, kemungkinan dapat membahayakan janinnya. Berdasarkan estimasi, Departemen Kesehatan (2003) mencatat jumlah pencemaran gas karbon monoksida dari sumber buatan diperkirakan mendekati 60 juta ton/tahun.

2.2.2. Sulfur Dioksida

Sulfur dioksida (SO₂) merupakan gas yang memiliki bau yang sangat tajam dan tidak mudah terbakar di udara. Sulfur dioksida terbentuk dari hasil pembakaran bahan – bahan yang mengandung sulfur didalamnya seperti batu bara, gas atau minyak. SO₂ bersifat larut dalam air dan dapat mengiritasi mata, kulit, selaput lendir dan sistem pernafasan serta pada konsesntrasi tinggi dapat menyebabkan kematian. Jika membentuk kabut (haze) dari reaksi fotokimia aerosol, SO₂, oksida nitrogen dan hidrokarbon di atmosfer. Senyawa sulfur dapat menurunkan jarak pandang, jika bereaksi dengan air hujan akan menimbulkan keasaman air hujan yang dapat menyebabkan asidifikasi sumber air serta penurunan unsur tanah, juga menyebabkan korosi logam dan bahan bangunan lain. Selain itu, Sulfur dioksida di atmosfer dapat berubah menjadi kabut asam sulfat (H₂SO₄) dan partikulat sulfat. (Ruhayat, 2009).

Akibat utama pencemaran gas sulfur oksida (SO₂) terhadap manusia adalah terjadinya iritasi pada sistem pernapasan. Beberapa penelitian menunjukkan bahwa iritasi tenggorokan terjadi pada konsentrasi SO₂ sebesar 5 ppm atau lebih. Bahkan pada beberapa individu yang sensitif, iritasi sudah terjadi pada

paparan 1-2 ppm. Bagi penderita yang mempunyai penyakit kronis pada sistem pernapasan dan kardiovaskular serta lanjut usia, dengan paparan yang rendah saja (0.2 ppm) sudah dapat menyebabkan iritasi tenggorokan. Lebih lengkap, pada tabel berikut ditunjukkan pengaruh SO_2 dalam berbagai kadar (ppm) terhadap kesehatan manusia.

Tabel 2. 1 Pengaruh Gas SO_2 Terhadap Manusia

| <i>Pengaruh Gas SO_2 Terhadap Manusia</i> | |
|---|--|
| <i>Kadar (ppm)</i> | <i>Dampaknya terhadap manusia</i> |
| 3 s.d. 5 | - Jumlah minimum yang dapat dideteksi baunya |
| 8 s.d. 12 | - jumlah minimum yang segera mengakibatkan iritasi tenggorokan |
| 20 | - Jumlah minimum yang mengakibatkan iritasi pada mata |
| | - Dapat menyebabkan batuk |
| | - Jumlah maksimum yang diperbolehkan untuk paparan yang lama |
| 50 s.d. 100 | - Jumlah maksimum yang dibolehkan untuk paparan yang singkat (+ 30 menit) |
| 400 s.d. 500 | - Sudah berbahaya walaupun dalam paparan yang singkat |

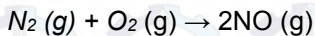
Sumber : Philip Kristanto, *Ekologi Industri*, Edisi Pertama cetakan pertama, 2002

Tidak hanya berdampak pada manusia, pencemaran SO_2 juga berdampak pada hewan dan tumbuhan. Sulfur dioksida SO_2 menyebabkan kerusakan pada pohon-pohon kayu dengan membentuk sulfit yang beracun pada sel daun pohon kayu tersebut (Hakiki, 2008). SO_2 dapat menurunkan jarak pandang, jika bereaksi dengan air hujan akan menimbulkan keasaman air hujan (Hujan Asam). Hujan asam terjadi ketika SO_2 bergabung dengan oksigen membentuk SO_3 yang bereaksi dengan butir-butir air

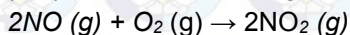
membentuk H_2SO_4 . hujan asam tersebut mampu mencemari biota ikan di danau atau di sungai (Hakiki, 2008)

2.2.3. Nitrogen Dioksida

Nitrogen dioksida merupakan polutan udara yang dihasilkan pada proses pembakaran gas nitrogen monoksida (NO). gas tersebut memiliki sifat tidak berwarna, yang pada konsentrasi tinggi juga dapat menimbulkan keracunan. Keberadaan gas nitrogen monoksida (NO) di udara disebabkan karena gas nitrogen ikut terbakar bersama dengan oksigen (O_2), yang terjadi pada suhu tinggi. Reaksinya adalah:



Pada saat kontak dengan udara, maka gas nitrogen monoksida (NO) akan membentuk gas NO_2 dengan reaksi sebagai berikut :



Sifat gas NO_2 merupakan gas yang beracun, berwarna merah cokelat, dan berbau seperti asam nitrat yang sangat menyengat dan merangsang. Pada suhu kamar kedua gas ini (Nitrogen dan Oksigen) hanya sedikit mempunyai kecenderungan untuk bereaksi satu sama lain. Pada suhu yang lebih tinggi (di atas $1210^\circ C$) keduanya dapat bereaksi membentuk nitrik oksida dalam jumlah tinggi sehingga mengakibatkan polusi udara. Dalam proses pembakaran, suhu yang digunakan biasanya mencapai $1210^\circ C - 1765^\circ C$. Jadi Nitrogen oksida (NO) dihasilkan dari buangan proses pembakaran dari transportasi dan akan segera teroksidasi di atmosfer membentuk NO_2 (Handriyanu, 2012)

Emisi antropogenik (aktivitas manusia) dari pembakaran minyak bumi, gas, dan bensin, dapat mengoksidasi nitrogen di atmosfer (N_2) membentuk NO yang ada di lingkungan. Sumber NO_2 di alam adalah dari proses mikrobiologi di dalam tanah. Di tanah, proses nitrifikasi dan denitrifikasi dapat melepaskan NO dan NO_2 ke atmosfer. Proses alami ini meningkat saat pupuk nitrogen ditambahkan untuk meningkatkan hasil pertanian. NO_x adalah katalis atmosfer yang berkaitan erat dengan pembentukan ozon dan radikal OH. NO_x berperan dalam mengatur proses oksidasi atmosferik dan siklus biogeokimia global (Firdaus, 2014)

Dampak yang ditimbulkan oleh pencemaran NO_2 sangat banyak. Menurut Handriyono (2012), Gas nitrogen oksida (NO_x)

ada dua macam , yakni gas nitrogen monoksida (NO) dan gas nitrogen dioksida (NO₂). Kedua macam gas tersebut mempunyai sifat yang berbeda dan keduanya sangat berbahaya bagi kesehatan. Gas NO yang mencemari udara secara visual sulit diamati karena gas tersebut tidak berwarna dan tidak berbau. Sedangkan gas NO₂ bila mencemari udara mudah diamati dari baunya yang sangat menyengat dan warnanya coklat kemerahan. Udara yang mengandung gas NO dalam batas normal relatif aman dan tidak berbahaya, kecuali jika gas NO berada dalam konsentrasi tinggi. Konsentrasi gas NO yang tinggi dapat menyebabkan gangguan pada system saraf yang mengakibatkan kejang-kejang. Bila keracunan ini terus berlanjut akan dapat menyebabkan kelumpuhan.

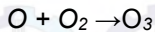
Gas NO akan menjadi lebih berbahaya apabila gas itu teroksidasi oleh oksigen sehingga menjadi gas NO₂. Keberadaan NO₂ lebih dari 1 ppm, dapat menyebabkan terbentuknya zat yang bersifat karsinogen atau penyebab terjadinya kanker dan Jika menghirup gas NO₂ dalam kadar 20 ppm akan dapat menyebabkan kematian. Studi sebelumnya mengindikasikan paparan NO_x terkait dengan efek kronis pada kesehatan paru-paru, terutama bagi populasi yang beresiko seperti anak kecil, orang tua dan penderita asma. NO_x merupakan polutan yang sering melewati standar kualitas udara ambien di daerah perkotaan (Driejana, 2006)

NO₂ berperan sebagai agen pengoksidasi yang kemungkinan merusak membran sel dan protein. Pada konsentrasi tinggi, saluran udara akan menyebabkan peradangan yang akut. Ditambah lagi, penyebaran dalam waktu-singkat berpengaruh terhadap peningkatan resiko infeksi saluran pernapasan. Meskipun banyak pengontrolan penyebaran yang dilakukan, fakta secara jelas mendefinisikan hubungan antara konsentrasi atau dosis dan umpan baliknya tidaklah cukup (Grineski, 2010).

2.2.4. Ozon

Ozon termasuk pencemar sekunder yang terbentuk di atmosfer dari reaksi fotokimia NO_x dan HC. Ozon bersifat

oksidator kuat, karena itu pencemaran oleh ozon troposferik dapat menyebabkan dampak yang merugikan bagi kesehatan manusia. Emisi gas buang berupa NO_x adalah senyawa pemicu (precursor) pembentukan ozon di lapisan atmosfer bawah (troposfer bawah pada ketinggian 0 – 2000m) terbentuk akibat adanya reaksi fotokimia pada senyawa oksida nitrogen (NO_x) dengan bantuan sinar matahari. Oleh karena itu potensi produksi ozon troposfer di daerah beriklim tropis seperti Indonesia sangat tinggi (Bappenas, 2008). Reaksi O₃ yang terbentuk dari polutan NO_x adalah sebagai berikut :



Ozon di muka bumi terbentuk oleh sinar ultraviolet yang menguraikan molekul O₃ membentuk unsur oksigen. Unsur oksigen ini bergabung dengan molekul yang tidak terurai dan membentuk O₃ kadangkala unsur oksigen akan bergabung dengan N₂ untuk membentuk nitrogen dioksida , jika bercampur dengan cahaya mampu membentuk ozon.

Ozon berwarna biru pucat, dan merupakan gas yang sangat beracun, serta berbau sangit. Ozon mendidih pada suhu - 199 °C dan mencair pada suhu -192,5 °C dan memiliki gravitasi 2.144. Ozon terbentuk ketika percikan listrik melintas dalam oksigen. Adanya ozon dapat dideteksi melalui aroma yang ditimbulkan oleh mesin mesin bertenaga listrik. Menurut Sudrajat (2005), ozon bersifat unik karena sebagai polutan sekunder dan juga bersifat sebagai oksidator. Sebagai polutan sekunder ozon mengalami pembentukan yang cepat dimana hanya membutuhkan waktu selama satu jam.

Keberadaan Ozon di tingkat bawah (troposfer) yang dihasilkan dari polutan-polutan hasil pembakaran dapat berdampak pada kesehatan manusia, dampak tersebut dapat berupa kenaikan tekanan darah, nyeri di dada, dan batuk. Dampak serius lainnya untuk kesehatan adalah memicu penyakit sesak nafas seperti bronkitis, emfisema, asma dan sakit jantung. Para ahli kesehatan mengemukakan bahwa manusia yang dalam kondisi sehat sekalipun dapat menjadi tersiksa jika secara konsisten terpapar oleh kadar ozon yang tinggi (Kerr, 2010).

Menurut Soedomo (2001), Ozon merupakan senyawa yang paling dominan dari oksidan fotokimia. Oksidan tersebut dapat merusak tanaman, karena setelah terpapar selama empat jam pada konsentrasi 100 mikrogram per meter kubik akan menimbulkan cacat pada daun

2.3. Partikulat Pencemar Udara

PM10 adalah partikel debu yang berdiameter kecil dari 10 mikron, yang mengandung karbon dan bahan bakar minyak. Sumber utama PM10 adalah di wilayah perkotaan tepatnya pada kendaraan bermesin diesel. Polutan partikel (PM10) mudah masuk ke tubuh manusia karena ukurannya yang sangat kecil.(Fardiaz, 1992).

Bahan partikulat PM10 menyertakan partikel partikel yang banyak menyebabkan efek buruk bagi kesehatan. Menurut US EPA (1992), ukuran partikel yang kecil, sangat mungkin masuk kedalam tubuh manusia terutama masuk ke dalam paru paru dan bisa menyebabkan gangguan pernafasan seperti Bronkitis, Asma atau bahkan kanker paru – paru.

2.4. Peramalan Deret Waktu (Time Series Forecasting)

2.4.1. Analisa Deret Waktu

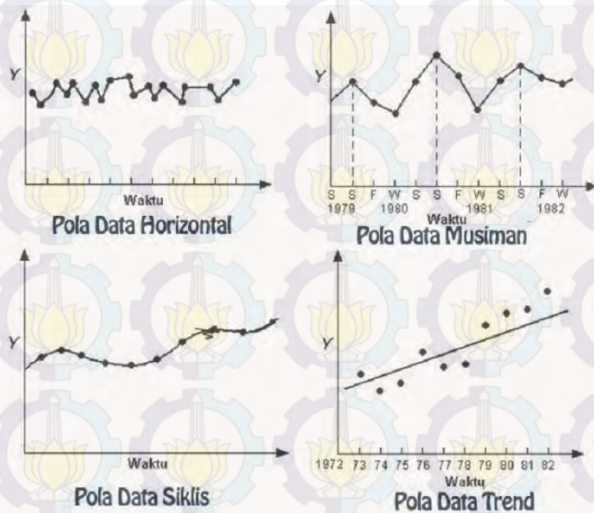
Deret Waktu atau Time series adalah serangkaian pengamatan terhadap variabel yang akan diamati secara berurutan dari waktu ke waktu dan dicatat berdasarkan urutan waktu kejadiannya (Wei, 2006). Analisis time series merupakan suatu metode peramalan untuk masa depan yang dilakukan berdasarkan nilai atau data masa lalu dari suatu variabel dan kesalahan (error) masa lalu. Tujuan dari metode peramalan time series ini adalah untuk menemukan pola data time series dan mengekstrapolasikan pola tersebut ke periode yang akan datang. Setiap pengamatan yang dilakukan dapat dinyatakan dalam bentuk variabel random Z_t , yang didapatkan berdasarkan indeks waktu tertentu t_i dengan $i = 1, 2, \dots, n$ sebagai urutan waktu

pengamatan, sehingga penulisan dari data time series adalah $Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}$.

2.4.2. Pengenalan Pola Data Deret Waktu

Pola data dalam peramalan deret waktu sangat beragam. Hal tersebut menjadi salah satu aspek penting dalam melakukan seleksi model peramalan yang sesuai untuk data deret waktu. Ada empat tipe pola data yang umum, yaitu : Horizontal, Trend, Seasonal dan Cyclical.

Pola Horizontal terjadi ketika data observasi berubah-ubah di sekitar rata-rata yang konstan. Pola trend terjadi ketika data observasi naik atau turun pada perluasan periode suatu waktu tertentu. Sedangkan pola siklikal ditandai dengan adanya fluktuasi data yang terjadi disekitar garis trend. Terakhir adalah pola data yang ditandai dengan adanya perubahan yang berulang secara otomatis dari waktu ke waktu, pola ini disebut sebagai pola musiman. Untuk lebih jelasnya empat tipe pola tersebut akan ditampilkan pada gambar 2.1

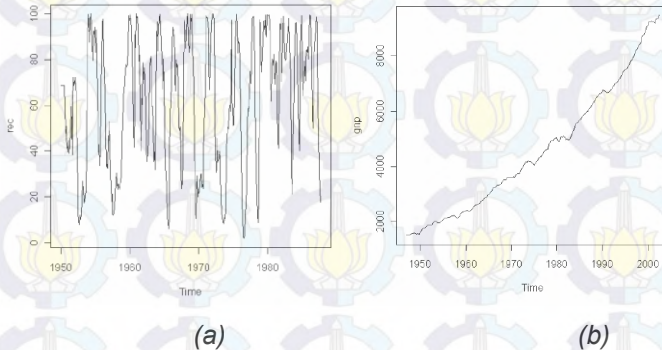


Gambar 2. 1 Pola Deret Waktu

2.4.3. Model ARIMA

Model Autoregressive Integrated Moving Average (ARIMA) merupakan salah satu model populer dengan menggabungkan model AR, MA, dan ARMA. Model AR, MA, dan ARMA diintegrasikan kedalam bentuk orde p , d , dan q . Prosedur standar yang digunakan dalam model ARIMA adalah prosedur Box-Jenkins. Prosedur Box-Jenkins terdiri dari 4 tahap, yaitu tahap identifikasi, estimasi, diagnostic check, dan peramalan (Bowerman dan O'Connell, 1993).

Model ARIMA dapat digunakan pada bentuk data yang bermacam-macam. Kasus-kasus yang dapat diprediksikan adalah data non-musiman/musiman yang stasioner, hingga data non-musiman/musiman yang tidak stasioner. Perbedaan model data diakibatkan pada perbedaan cara perlakuan data dalam prediksi di masa mendatang. Perlakuan signifikan dilakukan dengan transformasi box-cox dan differencing untuk model yang tidak stasioner menjadi stasioner (Suhartono, 2008).



Gambar 2. 2 Contoh data (a) Data Stasioner (b) Data Non-Stasioner

2.4.4. Model Seasonal Trend Decomposition Based On Loess (STL)

Seasonal Trend Decomposition using Loess (STL) adalah suatu bentuk algoritma yang dikembangkan untuk membantu mendekomposisikan sebuah time series menjadi 3 komponen yaitu trend, musiman dan remainder (sis).

Menurut Cleveland dkk Tujuan dibuatnya metode ini adalah untuk mengembangkan cara pendekomposisian agar pengguna dapat lebih mudah dalam memakai metode ini. Berikut adalah keistimewaan dari metode STL.

- 1. Desain yang sederhana dan mudah digunakan*
- 2. Fleksibel dalam menentukan jumlah variasi pada komponen trend dan seasonal*
- 3. Spesifikasi jumlah pengamatan tiap satu siklus pada komponen seasonal untuk bilangan bulat lebih besar dari Satu*
- 4. Mampu mendekomposisikan (menguraikan) data- data series yang tidak lengkap (hilang)*
- 5. Komponen seasonal dan komponen trend sangat kuat sehingga tidak akan mudah terdistorsi (,menyimpang)*
- 6. Penerapan pada komputer lebih mudah dan perhitungan lebih cepat, bahkan untuk time series yang cukup lama*

Definisi dari Loess itu sendiri adalah sebuah metode regresi non parametrik. Dimana regresi tersebut memiliki keunggulan, yaitu fleksibilitas yang tinggi karena data akan dengan sendirinya membentuk estimasi kurva yang tidak dipengaruhi oleh faktor subyektif (Outlier)

Secara Matematis, model STL dapat ditulis

$$Y_t = S_t + T_t + E_t \quad t = 1 \text{ sampai dengan } t = n$$

Dimana

- Y_t = nilai deret waktu sebenarnya pada periode t
 S_t = komponen musiman pada periode t
 T_t = komponen trend-siklus pada periode t
 E_t = komponen kesalahan pada periode t

Langkah langkah dalam dekomposisi STL

a. Inner Loop

Setiap iterasi pada inner loop terdiri dari sebuah pemulusan musiman yang memperbarui komponen musiman, diikuti sebuah pemulusan trend yang memperbarui komponen trend

b. Outer Loop

Outer loop digunakan untuk menghitung komponen sisa (random) berdasarkan trend-siklus dan komponen musiman yang dihasilkan.

Komponen sisa tersebut dilibatkan dalam estimasi deret keseluruhan melalui teknik kombinasi. Berdasarkan ekstraksi komponen sisa dari data pemisahan trend terekstrapolasi (\widehat{TR}_{T+1}) dan deret pemisahan komponen musiman terekstrapolasi (\widehat{SR}_{T+1}) Keduanya secara berurutan diperoleh dari menambahkan trend dengan sisa dan menambahkan musiman dengan sisa

Peramalan dengan Model STL

Peramalan komponen trend, komponen musiman, dan komponen trend-sisa, dan komponen musiman-sisa dilakukan sesuai dengan komponen tersebut. Dengan mengombinasikan keempat komponen tersebut dapat diperoleh estimasi peramalan untuk deret waktu keseluruhan sebagai berikut

$$\widehat{X}_{t+1} = \widehat{T}_{T+1} + \widehat{S}_{T+1} + \widehat{TR}_{T+1} + \widehat{SR}_{T+1} / 2$$

2.4.5. Stasioneritas Data

Stasioneritas berarti bahwa tidak terdapat perubahan yang drastis pada data. Fluktuasi data berada disekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut (Makridakis, 1995: 351). Data time series dikatakan stasioner jika rata-rata dan variansinya konstan, tidak ada unsur trend dalam data, dan tidak ada unsur musiman. Apabila data tidak stasioner, maka perlu dilakukan modifikasi untuk menghasilkan data yang stasioner. Salah satu cara yang umum dipakai adalah metode perbedaan (differencing).

Untuk menentukan apakah series stasioner, nonstasioner dapat dibantu dengan melihat plot dari series atau bentuk difference-nya. Proses differencing dapat dilakukan untuk beberapa periode sampai data stasioner, yaitu dengan cara mengurangi suatu data dengan data sebelumnya.

Stasioneritas dalam data time series ditunjukkan apabila rata-rata dan variansnya konstan setiap waktu. Untuk menentukan kestasioneran varians digunakan transformasi Box-Cox. Rumus umum dalam melakukan transformasi Box-Cox adalah sebagai berikut (Wei, 2006).

$$T(Z_t) = \frac{Z_t^\lambda - 1}{\lambda}, \text{ berlaku untuk } \lambda \neq 0 \quad (2.1)$$

Untuk melihat mengapa $\lambda = 0$ sesuai dengan logaritmik transformasi, dapat dituliskan sebagai berikut:

$$\lim_{\lambda \rightarrow 0} T(Z_t) = \lim_{\lambda \rightarrow 0} \frac{Z_t^\lambda - 1}{\lambda} = \ln(Z_t) \quad (2.2)$$

dimana λ merupakan parameter transformasi dari transformasi Box-Cox. Tabel 2.1 merupakan beberapa nilai λ yang biasanya digunakan pada transformasi Box-Cox.

Tabel 2.2 Transformasi Box-Cox

| Nilai Estimasi λ | Transformasi |
|--------------------------|--------------------------------|
| -1 | $1/Z_t$ |
| -0,5 | $1/\sqrt{Z_t}$ |
| 0 | $\ln Z_t$ |
| 0,5 | $\sqrt{Z_t}$ |
| 1 | Z_t (tidak ada transformasi) |

Selanjutnya dilakukan identifikasi kestasioneran data terhadap mean. Identifikasi kestasioneran terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan menggunakan plot ACF. Data time series bersifat stasioner terhadap mean jika plot time series berfluktuasi disekitar nilai rata-rata yang konstan. Apabila data time series dinyatakan tidak stasioner dalam mean, maka langkah selanjutnya adalah melakukan differencing. Rumus differencing adalah sebagai berikut (Wei, 2006).

$$Y_t = (1 - B)^d Z_t \quad (2.3)$$

2.4.6. Fungsi Autokolerasi dan Autokorelasi Parsial

Autokorelasi atau Autocorrelation Function (ACF) adalah koefisien yang menunjukkan keeratan hubungan antara nilai variabel yang sama tetapi pada waktu yang berbeda. Menurut Gujarati (2003) Uji autokorelasi ini dapat didefinisikan sebagai korelasi antara anggota serangkaian observasi yang diurutkan menurut waktu (seperti dalam data time series). Fungsi uji autokorelasi adalah untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik autokorelasi yaitu korelasi yang

terjadi antara residual pada satu pengamatan dengan pengamatan lain pada model regresi.

Korelasi parsial adalah pengukuran hubungan antara dua variabel, dengan mengontrol atau menyesuaikan efek dari satu atau lebih variabel lain. Fungsi parsial autokorelasi atau Partial Autocorrelation Function (PACF) merupakan koefisien autokorelasi parsial yang mengukur derajat hubungan antara nilai-nilai sekarang dengan nilai-nilai sebelumnya (untuk time lag tertentu), sedangkan pengaruh nilai variabel time lag yang lain dianggap konstan. Fungsi autokorelasi parsial adalah korelasi antara Z_t dan Z_{t+k} setelah pengaruh dari variabel pengganggu $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$ dihilangkan.

2.4.7. Indikator Keakuratan Peramalan

Pengukuran keakuratan peramalan model perlu dilakukan untuk mengetahui tingkat kemampuan prediksi suatu model terhadap data-data yang telah diketahui. Kesesuaian model dapat diketahui dengan mempelajari ketepatan nilai ramalan pada prediksi runtun waktu. Ukuran yang biasa digunakan untuk mengevaluasi ketepatan ramalan adalah menggunakan ukuran statistik standar. Ukuran tersebut terdiri dari MAPE (Mean Absolute Percentage Error). Formula yang digunakan pada ukuran-ukuran tersebut adalah :

$e_i = X_i - F_i$ = Kesalahan ramalan periode i .

X_i = Data aktual periode i .

F_i = Data hasil ramalan periode i .

n = Banyaknya Jumlah Observasi

- PE (Percentage Error) adalah presentase kesalahan dari nilai aktual X_i dengan hasil ramalan F_i

$$|PE| = \frac{|X_i - F_i|}{X_i}$$

- *MAPE (Mean Absolute Percentage Error) adalah rata-rata kesalahan, namun memberikan nilai absolute pada selisih aktual dengan nilai hasil peramalan. MAPE merupakan indikator yang biasa digunakan untuk menunjukan performance atau keakuratan pada hasil proses ramalan*

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

- *Batasan nilai untuk indikator MAPE menurut Maridakis dan Hibon tahun 1990 adalah :*

$$0 \leq MAPE \leq +\infty$$

Evaluasi kesesuaian model juga ditentukan terhadap pola grafik yang dinilai berdasarkan koefisien determinasi atau yang biasa disimbolkan dengan R^2 . Koefisien determinasi menunjukkan kecenderungan pola grafik antara nilai-nilai estimasi prediksi dan nilai-nilai aktual. Nilai R^2 yang semakin mendekati angka 1 menunjukkan pola grafik prediksi dan aktual yang sama. (Suwardiwijaya dkk., 2009)

2.5. Software Statistik R

Software statistik berbentuk freeware merupakan solusi dari lisensi mahal software statistik komersil. R merupakan kelompok paket software statistik untuk keperluan analisis data yang berbentuk open source atau freeware. Paket statistik R bersifat multiplatform yang tersedia untuk berbagai macam sistem operasi.

R adalah suatu kesatuan software yang terintegrasi dengan beberapa fasilitas untuk manipulasi, perhitungan dan penampilan grafik yang handal. R berbasis pada bahasa pemrograman S, yang dikembangkan oleh AT&T Bell Laboratories (sekarang Lucent Technologies) pada akhir tahun '70 an.

R merupakan versi gratis dari bahasa S dari software (berbayar) yang sejenis yakni S-PLUS yang banyak digunakan para peneliti dan akademisi dalam melakukan kegiatan ilmiahnya.

Menurut Ihaka dan Gentleman (1996), ada beberapa alasan mengapa menggunakan R, diantaranya adalah:

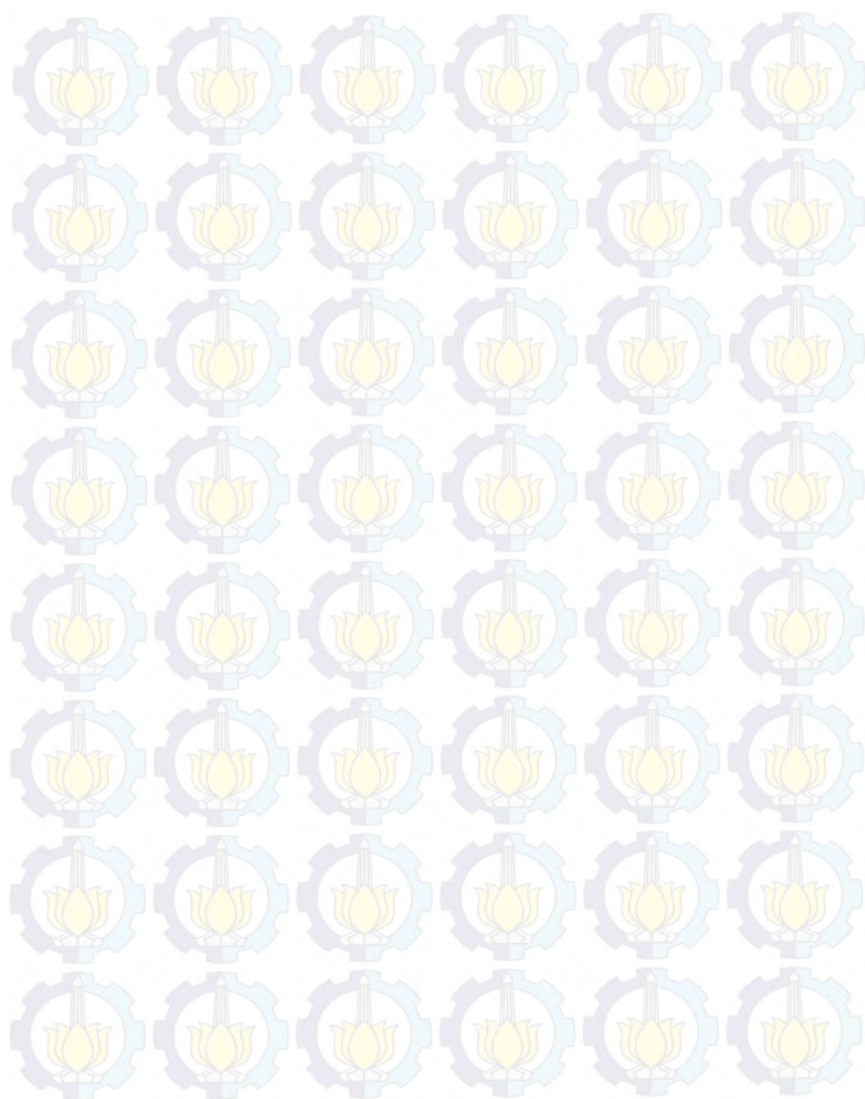
- 1) *Software R* sangat serbaguna dikarenakan berorientasi obyek dan memiliki banyak *library* yang sangat bermanfaat yang dikembangkan oleh kontributor. Pengguna bebas menambah dan mengurangi *library* tergantung kebutuhan.
- 2) Efektif dalam pengelolaan data dan fasilitas penyimpanan. Ukuran file yang disimpan jauh lebih kecil dibanding software lainnya.
- 3) Lengkap dan terdiri dari koleksi tools statistik yang terintegrasi untuk analisis data, diantaranya, mulai statistik deskriptif, fungsi probabilitas, berbagai macam uji statistik, hingga *time series*.
- 4) Tampilan grafik yang menarik dan fleksibel ataupun *costumized*
- 5) Dapat dikembangkan sesuai keperluan dan kebutuhan dan sifatnya yang terbuka, setiap orang dapat menambahkan fitur-fitur tambahan dalam bentuk paket ke dalam software R

Software R dapat diperoleh pada CRAN-archive yaitu The Comprehensive R Archive Network di laman <http://cran.r-project.org>. Kemampuan utama dari R terletak pada Add-on packages/library yang berfungsi sebagai kumpulan perintah (Suhartono, 2008).

2.5.1. Paket Library

Software R menyediakan para pengguna R untuk mengembangkan model statistika. Hasil pengembangan-pengembangan model dikumpulkan dalam satu paket *library* dan harus dipasang satu per satu sesuai dengan kebutuhan analisis.

Paket *Library* 'forecast' memuat fungsi statistika yaitu *tslm* (*Linear model with seasonal components*) dan juga *ARIMA model*. *Library* 'fracdiff' sebagai paket fungsi fractional difference untuk *ARIMA(p,d,q)* dan 'MSwM' sebagai fungsi peubah markov model *linear*.



BAB 3 METODE PENELITIAN

3.1. Gambaran Umum

Penelitian ini tentang Prediksi konsentrasi kualitas udara yang menggunakan dua model, yaitu Seasonal Decomposition Based on Loess (STL) dan ARIMA. Metode penelitian ini disusun secara sistematis sehingga proses penelitian dari awal hingga akhir dapat berjalan dengan lancar. Tujuan dari metode penelitian ini adalah:

- 1. Alat kontrol penelitian untuk mengetahui pelaksanaan penelitian dan pembuatan laporan tugas akhir dapat berjalan secara sistematis.*
- 2. Meminimalkan terjadinya kesalahan selama penelitian berlangsung.*
- 3. Memberikan batasan lingkup penelitian yang jelas, sehingga tujuan penelitian dapat tercapai dengan memberikan gambaran hal-hal yang berkaitan terhadap penelitian.*

Penelitian ini akan dilakukan dengan beberapa tahap, yaitu perumusan gagasan penelitian, studi literatur, penentuan skenario model, input data in-sample, simulasi model, hasil dan pembahasan, serta ditutup dengan kesimpulan dan saran.

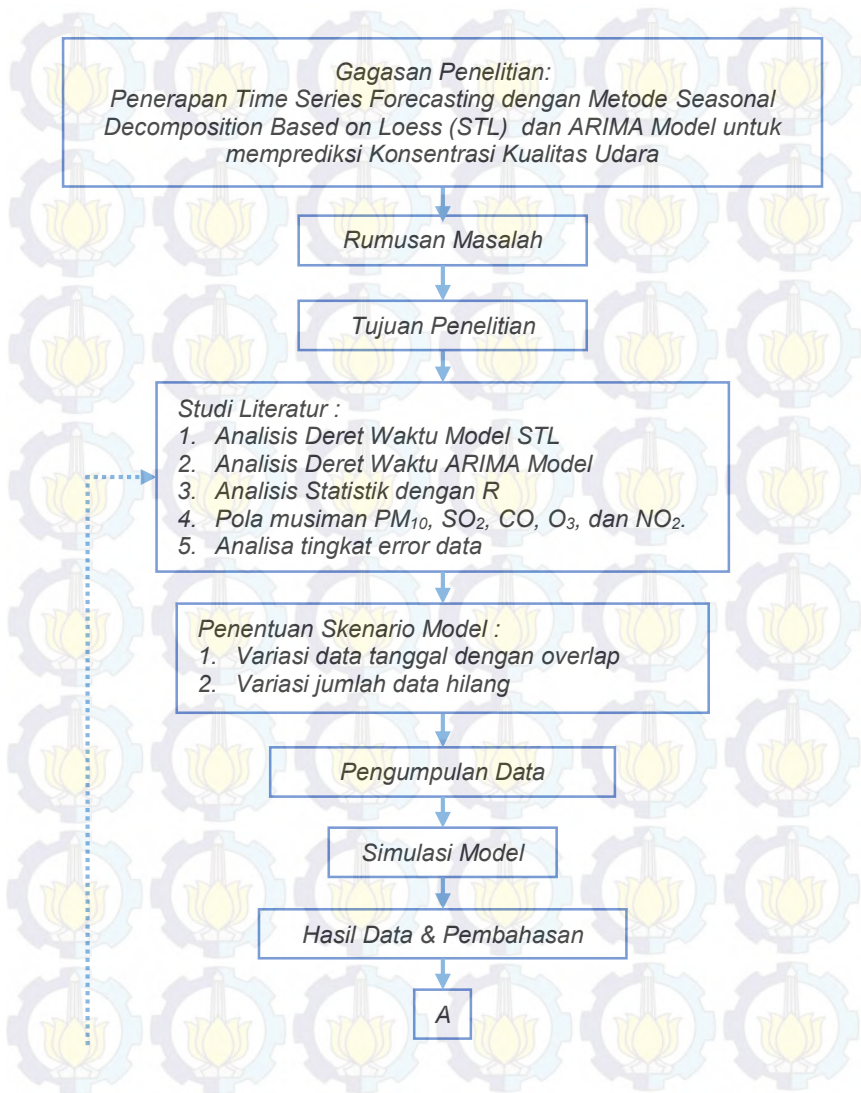
3.2. Lokasi dan Lingkungan Sekitar SUF

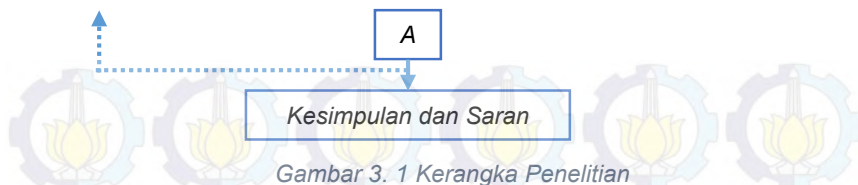
Penelitian prediksi konsentrasi udara ini menggunakan data stasiun SUF 7 Kebonsari Surabaya yang terletak pada koordinat 7°19'41.30" LS dan 112°42'49.00" BT. SUF 7 berada di wilayah kecamatan Gayungan yang memiliki luas wilayah sebesar 6,07 km². SUF 7 Kebonsari terletak di jalan Gayung Kebonsari Raya. Lokasi SUF 7 mewakili daerah perkampungan lama dan baru dengan area militer yang berada di utara dan area pergudangan/industri dan perdagangan/jasa di wilayah timur stasiun pemantau

Stasiun pemantau SUF 7 Kebonsari terletak di sebelah barat jalan tol Surabaya – Gresik dan di sebelah timur sungai Kali Surabaya. Permukiman sekitar SUF berada pada angka 6563 jiwa/km² per tahun 2014 (BPS Surabaya, 2014).

3.3. Kerangka Penelitian

Kerangka penelitian merupakan bagan alir penelitian untuk memberikan tahapan sistematis terhadap kegiatan penelitian. Kerangka penelitian digunakan menjadi acuan kerja dari awal hingga berakhirnya penelitian. Kerangka penelitian selengkapnya dapat dilihat pada gambar 3.1





Gambar 3. 1 Kerangka Penelitian

3.4. Studi Literatur

Studi literatur dilakukan untuk menunjang dasar ilmiah penelitian ini agar dapat memberikan wawasan untuk penelitian dalam memprediksi ini. Penelitian-penelitian terdahulu juga dijadikan literatur sehingga kesalahan di penelitian dapat diminimisasi. Studi literatur yang harus dicari untuk menunjang penelitian ini, berkaitan dengan:

1. Analisis Deret Waktu ARIMA Model
2. Analisis Deret Waktu STL Model
3. Analisis Statistik dengan R Studio
4. Pola musiman PM_{10} , SO_2 , CO , O_3 , dan NO_2 .
5. Analisa tingkat error data

Studi literatur ini dilakukan sejak awal penelitian hingga penarikan kesimpulan. Adanya studi literatur juga untuk memberikan dasar yang kuat terhadap keseluruhan pada penelitian sehingga penarikan kesimpulan yang didapat dari penelitian dapat dipertanggungjawabkan.

3.5. Penentuan Skenario Model

Skenario model dilakukan dengan menentukan kondisi terbaik untuk hari ke-15 sebagai data acuan terhadap hasil prediksi. Skenario model dibagi menjadi dua bagian yaitu kondisi data aktual dimana terdapat overlapping data dan kondisi data buatan dengan manipulasi jumlah data hilang. Overlapping data adalah data dimana dalam setiap percobaannya terdapat hari yang juga digunakan untuk percobaan selanjutnya.



Gambar 3. 2 Gambaran Umum Overlapping Data

Penggunaan overlapping data bertujuan untuk menemukan tingkat konsistensi model terhadap data historis di tiap parameter. Overlapping juga berfungsi untuk membuat data konsentrasi menjadi homogen (tidak terpaut nilai yang tinggi antara arah angin, suhu, kelembapan, dan lain-lain). Data yang digunakan untuk percobaan overlapping data ditunjukkan pada tabel 3.1.

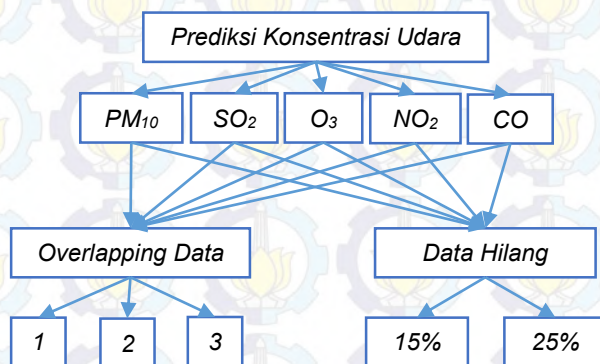
Tabel 3.1 Tanggal Data di Tiap Parameter

| Parameter | Tanggal <i>in-sample</i> | | |
|------------------|--------------------------|----------------|-------------------|
| PM ₁₀ | 43 Data Hilang | 37 Data Hilang | 38 Data Hilang |
| | 9 – 23 Jul 14 | 12 – 26 Jul 14 | 15 – 29 Jul 14 |
| SO ₂ | 25 Data Hilang | 25 Data Hilang | 23 Data Hilang |
| | 4 – 18 Aug 14 | 7 – 21 Aug 14 | 10 – 24 Aug 14 |
| CO | 14 Data Hilang | 14 Data Hilang | 14 Data Hilang |
| | 12 – 26 Jul 14 | 15 – 29 Jul 14 | 18 Jul – 1 Aug 14 |
| O ₃ | 14 Data Hilang | 14 Data Hilang | 14 Data Hilang |
| | 13 – 27 Jul 14 | 16 – 30 Jul 14 | 19 Jul – 2 Aug 14 |
| NO ₂ | 19 Data Hilang | 18 Data Hilang | 18 Data Hilang |
| | 29 Jul – 11 Aug 14 | 1 – 15 Aug 14 | 4 – 18 Aug 14 |

Overlapping data yang dilakukan sebanyak 3 kali pada jarak hari yang sama yaitu sebanyak 3 hari. Tanggal yang dipilih berdasarkan jadwal musim kemarau (April – September) dan hasil pencatatan alat pantau terbaik. Pencatatan alat pantau terbaik didasarkan atas jumlah data hilang pada tanggal data out-sample (hari ke-15).

Variasi dalam jumlah data hilang bertujuan untuk melihat tingkat kemampuan model terhadap data hilang yang berlebih. Hal ini dikondisikan terhadap SUF 7 yang memiliki variasi jumlah data hilang pada tiap harinya. Data hilang pada SUF 7 maksimal berada pada angka 12 data hilang di tiap harinya (parameter O_3). Variasi data hilang sebanyak 15% (50 data) dan 25% (84 data). Pengambilan data hilang sebanyak 50 data diharapkan dapat memenuhi gambaran lapangan tentang data hilang terhadap konsistensi model forecasting. Bentuk skenario ini merupakan skala kecil peramalan yang akan digunakan nantinya jika diaplikasikan di lapangan.

Berikut ini gambar kerangka bagan skenario model yang dilakukan:



Gambar 3. 3 Skenario Model Time Series Forecasting

3.6. Pengumpulan Data

Pengumpulan data adalah awal dari simulasi model yang dimulai dari data sekunder awal hingga data siap untuk digunakan pada simulasi model. Data yang digunakan pada overlapping data menggunakan data SUF 7 Kebonsari yang didapatkan pada BLH Kota Surabaya.

Tahapan-tahapan manajemen data adalah konversi data harian menjadi data 15 harian, konversi data per 30 menit menjadi 1 jam, dan konversi ekstensi data dari .xls (excel spreadsheet)

menjadi .csv (comma delimited). Tahapan-tahapan manajemen data dibentuk kedalam bagan berikut:

Data sekunder awal merupakan data yang masih berupa harian dengan frekuensi 30 menit di tiap hari. Data olahan berikutnya bertujuan untuk mendapatkan hasil berupa data dengan frekuensi 60 menit dan panjang data sebanyak 15 hari (ekstensi telah dirubah menjadi .csv).



Sistem Pemantauan Kualitas Udara Kota Surabaya
Laporan harian kualitas udara di stasiun pemantauan
Daily Air Quality Report for Monitoring Station
Tanggal / Date: 01.01.2015
Kota / City: Surabaya
Stasiun / Station: SUF7 (Kebonsari)

Laporan disiapkan di / Report prepared at: Data Center Surabaya
Laporan berdasarkan data setengah jam rata-rata / Report based on half-hourly values

| Waktu / Time GMT+7 (WIB) | PM10 ug/m3 | SO2 ug/m3 | CO mg/m3 | O3 ug/m3 | NO2 ug/m3 | NO ug/m3 |
|-----------------------------|---------------|--------------|-------------|-------------|--------------|-------------|
| 00:30 | 58.00 | 4.40 | 1.41 | 33.26 | 0.00 | 1.98 |
| 01:00 | 59.75 | 3.99 | 1.42 | 33.24 | 0.00 | 0.80 |
| 01:30 | 61.00 | 4.13 | 1.80 | 29.19 | 0.00 | 7.69 |
| 02:00 | 60.38 | 4.27 | 1.64 | 27.24 | 0.00 | 13.11 |
| 02:30 | 60.00 | 4.49 | 1.63 | 29.32 | 0.00 | 7.36 |
| 03:00 | 57.43 | 4.26 | 1.32 | 31.46 | 0.00 | 0.00 |
| 03:30 | 56.00 | 4.41 | 1.20 | 32.14 | 0.00 | 0.00 |
| 04:00 | 56.67 | 4.58 | 0.98 | 33.16 | 0.00 | 0.00 |
| 04:30 | 57.00 | 5.02 | 1.26 | 27.68 | 0.00 | 1.52 |
| 05:00 | 54.85 | 4.91 | 1.52 | 26.64 | 0.00 | 4.72 |
| 05:30 | 54.00 | 4.90 | 1.54 | 26.93 | 0.00 | 2.53 |
| 06:00 | 54.00 | 5.16 | 1.71 | 26.41 | 0.00 | 6.63 |
| 06:30 | 54.00 | 5.75 | 1.49 | 30.39 | 0.00 | 5.42 |
| 07:00 | 54.70 | 5.64 | 1.39 | 33.98 | 0.00 | 3.83 |
| 07:30 | 55.00 | 5.57 | 1.25 | 39.19 | 0.00 | 1.47 |
| 08:00 | 56.63 | 5.00 | 1.19 | 46.50 | 0.00 | 0.23 |
| 08:30 | 57.00 | 4.41 | 1.25 | 53.17 | 0.00 | 0.77 |
| 09:00 | 51.95 | 4.32 | 1.33 | 58.29 | 0.00 | 0.96 |
| 09:30 | 51.00 | 4.11 | 1.13 | 56.36 | 0.00 | 1.23 |
| 10:00 | 44.00 | 4.10 | 1.03 | 61.44 | 0.00 | 0.00 |
| 10:30 | 43.00 | 3.87 | 1.04 | 64.82 | 0.00 | 0.00 |
| 11:00 | 37.50 | 4.01 | 1.10 | 71.70 | 0.00 | 0.00 |
| 11:30 | 37.00 | 4.39 | 1.04 | 88.87 | 0.00 | 0.00 |
| 12:00 | 36.07 | 4.74 | 1.18 | 108.18 | 0.00 | 0.00 |
| 12:30 | 36.00 | 5.22 | 1.40 | 113.37 | 0.00 | 0.00 |
| 13:00 | 38.92 | 5.31 | 1.10 | 84.87 | 0.00 | 0.00 |
| 13:30 | 38.98 | 4.33 | 0.83 | 73.92 | 0.00 | 0.00 |
| 14:00 | 37.00 | 4.79 | 0.94 | 82.28 | 0.00 | 0.00 |
| 14:30 | 36.97 | 4.79 | 1.02 | 86.96 | 0.00 | 0.00 |
| 15:00 | 36.00 | 4.76 | 1.06 | 78.19 | 0.00 | 0.00 |
| 15:30 | 35.73 | 4.30 | 1.03 | 87.74 | 0.00 | 0.00 |
| 16:00 | 32.00 | 4.07 | 1.12 | 88.72 | 0.00 | 0.00 |
| 16:30 | 32.00 | 3.79 | 0.94 | 55.29 | 0.00 | 3.11 |
| 17:00 | 32.00 | 3.64 | 0.90 | 46.60 | 0.00 | 2.58 |
| 17:30 | 33.33 | 4.20 | 1.29 | 36.39 | 0.00 | 15.94 |
| 18:00 | 42.00 | 3.92 | 1.32 | 33.63 | 2.30 | 10.72 |
| 18:30 | 41.65 | 3.75 | 1.66 | 31.58 | 11.76 | 25.93 |
| 19:00 | 40.00 | 3.82 | 1.84 | 33.11 | 9.79 | 13.79 |
| 19:30 | 39.79 | 3.73 | 1.36 | 36.80 | 1.41 | 3.95 |
| 20:00 | 39.00 | 3.61 | 1.18 | 33.90 | 0.00 | 4.11 |
| 20:30 | 38.07 | 3.79 | 1.43 | 31.72 | 3.20 | 13.89 |
| 21:00 | 35.00 | 3.76 | 1.66 | 29.70 | 7.86 | 21.25 |
| 21:30 | 35.55 | 4.03 | 1.82 | 31.46 | 0.00 | 70.03 |
| 22:00 | 37.00 | 3.94 | 1.45 | 31.26 | 3.02 | 40.52 |
| 22:30 | 35.77 | 4.83 | 1.35 | 31.26 | --- | --- |
| 23:00 | --- | --- | --- | --- | --- | --- |
| 23:30 | 34.00 | 8.92 | 1.28 | --- | 0.19 | 38.68 |
| 24:00 | 36.00 | 4.83 | 1.37 | 30.21 | 0.00 | 45.36 |
| Minimum / Minimum | 32.00 | 3.61 | 0.83 | 26.41 | 0.00 | 0.00 |
| Nilai Rata-rata / Mean | 44.89 | 4.52 | 1.30 | 49.08 | 0.86 | 8.05 |
| Maksimum / Maximum | 61.00 | 8.92 | 1.84 | 113.37 | 11.76 | 70.03 |

| O3 |
|------------|
| 1 33.254 |
| 2 28.211 |
| 3 30.3935 |
| 4 32.651 |
| 5 27.163 |
| 6 26.6685 |
| 7 32.1855 |
| 8 42.8475 |
| 9 55.728 |
| 10 58.9025 |
| 11 68.2615 |
| 12 98.5255 |
| 13 99.1195 |
| 14 78.1015 |
| 15 82.574 |
| 16 88.2275 |
| 17 50.945 |
| 18 35.0095 |
| 19 32.345 |
| 20 34.8485 |
| 21 30.7075 |
| 22 31.364 |
| 23 NA |
| 24 NA |

Gambar 3. 4 Gambaran Pengolahan data O3

Data 15 harian dibagi menjadi 2 bagian yaitu data training (in-sample) dan data verifikasi (out-sample). Data training yang digunakan pada penelitian ini sebanyak 14 hari dengan 24 data di setiap harinya. Data in-sample ini merupakan data yang digunakan sebagai data runtun waktu (data historis) parameter dan dijadikan sebagai acuan untuk memprediksikan nilai-nilai konsentrasi di hari ke 15 atau yang disebut data out-sample. Data out-sample berisikan sebanyak 23 data dengan penghapusan nilai di pukul 23:30 hingga 24:00 dikarenakan data yang hilang pada waktu tersebut dan data verifikasi tidak dapat diganti dengan nilai lain.

3.7. Simulasi Model

Simulasi model ini dilakukan dengan menggunakan software statistik R dan bertujuan untuk menentukan hubungan terhadap skenario yang telah dibuat sehingga dapat ditentukan model manakah yang paling konsisten terhadap parameter uji, serta konsistensi model terhadap jumlah data hilang.

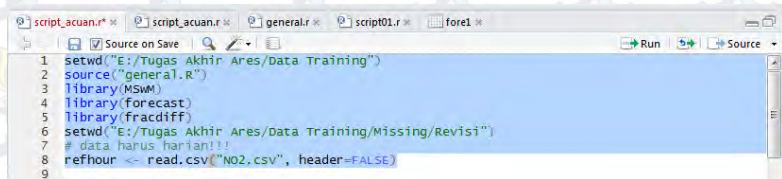
Simulasi model ini akan dilaksanakan dalam beberapa tahap. Simulasi model berisi tahapan proses input data, olah data, dan output data ini akan dilakukan sebagai berikut:

1. Proses input data

Proses input data yang pertama adalah Setwd atau set working directory adalah sebuah perintah yang bertujuan untuk mengatur letak penempatan data (working directory). Pada penelitian ini menggunakan dua kali perintah setwd. Yang pertama untuk penempatan data general R dan yang ke dua untuk penempatan data konsentrasi zat pencemar (.csv)

Paket library adalah sekumpulan fungsi atau rumus R yang digunakan untuk simulasi model. Package library yang digunakan pada prediksi deret waktu ini adalah forecast (berisi perintah auto.arima dan tslm), MSwM (berisi perintah peubah markov), dan fraccdiff (untuk melakukan difference pada ARIMA model). Data yang telah diolah hingga menjadi ekstensi csv merupakan bahan utama untuk melakukan input data. Input data dilakukan dengan

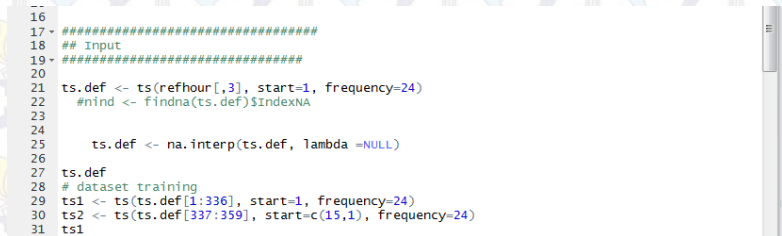
membe rikan lokasi direktori kerja (working directory) dan memanggil nama data hasil olahan berekstensi csv



```
1 setwd("E:/Tugas Akhir Ares/Data Training")
2 source("general.R")
3 library(MSWM)
4 library(forecast)
5 library(fracdiff)
6 setwd("E:/Tugas Akhir Ares/Data Training/Missing/Revisi")
7 # data harus harian!!!
8 refhour <- read.csv("NO2.csv", header=FALSE)
9
```

Gambar 3. 5 Script untuk memanggil Package Library dan Data Sampel

Data yang telah dimasukkan kedalam program R kemudian dibagi menjadi data *in-sample* dan *out-sample*. Data *in-sample* diberikan nama *ts1* (untuk data deret waktu 1) dan *out-sample* diberikan nama *ts2* (untuk data deret waktu 2). Data *ts1* menjadi dasar prediksi dan data *ts2* menjadi dasar penentuan kemampuan model selanjutnya.



```
16
17 - #####
18 ## Input
19 - #####
20
21 ts.def <- ts(refhour[,3], start=1, frequency=24)
22 #mind <- findna(ts.def)$indexNA
23
24
25 ts.def <- na.interp(ts.def, lambda =NULL)
26
27 ts.def
28 # dataset training
29 ts1 <- ts(ts.def[1:336], start=1, frequency=24)
30 ts2 <- ts(ts.def[337:359], start=c(15,1), frequency=24)
31 ts1
```

Gambar 3. 6 Proses Pembagian Data

2. Proses olah data

Proses olah data dilakukan dengan menjalankan perintah simulasi model yang akan digunakan. Linear model menjadi "**model 1**" dan ARIMA model menjadi "**model 2**". Pemberian nama model untuk mempermudah dalam penyebutan selanjutnya. Prediksi konsentrasi udara untuk model linear menggunakan perintah *tslm* sedangkan model ARIMA menggunakan perintah *auto.arima*.

Perintah ETS akan secara otomatis melakukan perhitungan dengan Model STL pada data *in-sample* dan diproyeksikan

terhadap data out-sample dengan memperhitungkan pola musiman. Perintah `auto.arima` adalah perintah yang secara otomatis menentukan orde terbaik untuk prediksi ARIMA dengan memberikan orde musiman, differencing data, dan transformasi Box-Cox dengan pemberian parameter transformasi (λ).

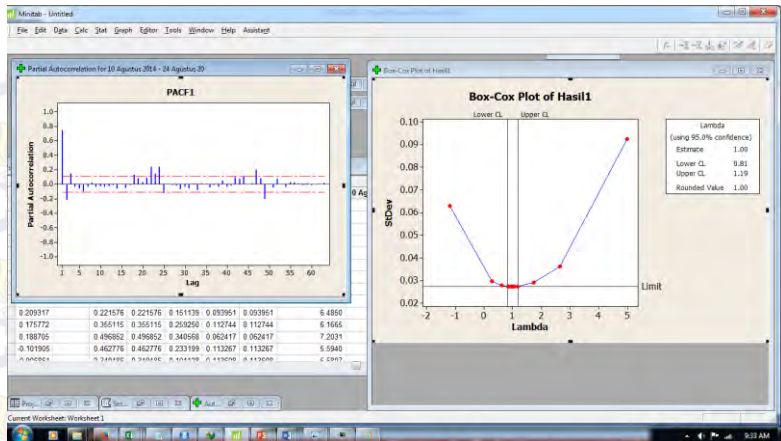
```

38 - #####
39 ## Model 1 - ets with stl data
40 - #####
41
42
43 ## ngefit ets using stl data
44 model1fit2 <- stl(ts1, s.window="periodic")
45
46 ## forecasting
47 model1fore <- forecast(model1fit2, method=c("ets"), h=23)
48
49 fit <- stl(ts1, s.window="periodic")
50 plot(forecast(fit))
51
52 - #####
53 ## Model 2 - ARIMA
54 - #####
55
56 - #####
57 ## TRANSFORMASI BOX COX
58 - #####
59
60 ## milih parameter transformasi boxcox
61 BoxCox.lambda(ts1, method=c("guerrero"), lower=-1, upper=2)
62 #BoxCox.lambda(ts1, method=c("loglik"), lower=-5, upper=5)
63
64 ## ngefit ARIMA, check order from above
65 model2fit2 <- auto.arima(ts1, lambda=1)
66

```

Gambar 3. 7 Perintah STL dan `auto.arima`

Transformasi Box Cox dipermudah dengan bantuan perangkat lunak minitab. Dengan software tersebut dapat dengan mudah mencari stasioner data dan menentukan lambda yang tepat untuk transformasi Box-Cox tersebut.



Gambar 3. 8 contoh pengolahan data dengan minitab

3. Proses output data

Data keluaran yang dihasilkan oleh software statistik R berupa plot grafik dan kesesuaian model. Kriteria kesesuaian model berbeda di tiap parameter. Kriteria berisikan ME, RMSE, R2, dan CORR. Kriteria evaluasi model untuk menentukan kemampuan model. Output berupa grafik ditampilkan dengan membandingkan kedua model. Berikut adalah kriteria standart pada untuk menyeleksi model terbaik

Hasil simulasi model yang memiliki nilai melebihi angka kriteria menunjukkan model yang tidak mampu untuk digunakan prediksi.

3.8. Hasil dan Pembahasan

Analisa hasil dan pembahasan akan dilakukan antara kemampuan model dalam memprediksikan konsentrasi dan hubungannya antara model lain dalam konsistensi dan juga dengan perbedaan jumlah data hilang.

- 1. Hubungan antara Koefisien Determinan (R^2), Corelation (Korelasi) serta Mean Absolute Percentage Error (MAPE)*
- 2. Hubungan antara konsistensi model antara model 1 dan 2.*
- 3. Hubungan antara jumlah data hilang dengan kemampuan model.*

Seelah analisa antara hubungan telah dilakukan maka akan dicari model terbaik yang dapat digunakan untuk memprediksikan konsentrasi udara di wilayah SUF 7 Kebonsari Surabaya.

3.9. Kesimpulan dan Saran

Kesimpulan yang akan ditarik dapat berupa:

- 1. Kemampuan model prediksi terhadap masing-masing parameter.*
- 2. Konsistensi model prediksi terhadap dampak perbedaan hari.*
- 3. Konsistensi model prediksi terhadap dampak perbedaan jumlah data hilang dengan hari yang sama.*

Saran yang akan ditarik berisikan tentang perbaikan-perbaikan dan juga variabel-variabel yang belum disampaikan di penelitian ini.

BAB 4

ANALISA DAN PEMBAHASAN

4.1. Performa Model STL dan Model ARIMA terhadap Perbedaan Data *In Sample*

Performa model Seasonal Decomposition based on Loess (STL) dan ARIMA ditentukan dengan kemampuan tiap model dalam mencari data yang hilang pada waktu yang berbeda. Pada penelitian ini pengujian performa model menggunakan tiga data training yang overlapping sebanyak 3 hari (terjadi pengulangan data training). dengan tujuan untuk melihat konsistensi performa dari tiap tiap model.

4.1.1. SO₂

1. Data Tanggal 04 Agustus 2014 - 18 Agustus 2014

Pengujian model menggunakan data in sample pada tanggal 4 Agustus 2014 hingga tanggal 17 Agustus 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 18 Agustus 2014

4.

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R² dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R.

Tabel 4. 1 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|--------|-----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 10.720 | 8.273 | 2.447 | 5.990 | 23% |
| 2 | 15.407 | 10.667 | 4.740 | 22.472 | 31% |
| 3 | 21.262 | 12.273 | 8.988 | 80.785 | 42% |
| 4 | 35.495 | 10.386 | 25.109 | 630.468 | 71% |

| | | | | | |
|------------------|----------------|----------------|----------------|-----------------|------------|
| 5 | 44.300 | 8.856 | 35.444 | 1256.291 | 80% |
| 6 | 31.515 | 10.387 | 21.128 | 446.408 | 67% |
| 7 | 33.757 | 13.501 | 20.256 | 410.301 | 60% |
| 8 | 32.524 | 17.265 | 15.259 | 232.848 | 47% |
| 9 | 18.975 | 9.800 | 9.176 | 84.197 | 48% |
| 10 | 11.434 | 5.993 | 5.441 | 29.606 | 48% |
| 11 | 11.359 | 5.472 | 5.886 | 34.646 | 52% |
| 12 | 10.948 | 5.248 | 5.699 | 32.480 | 52% |
| 13 | 11.348 | 5.181 | 6.167 | 38.032 | 54% |
| 14 | 9.902 | 5.189 | 4.713 | 22.211 | 48% |
| 15 | 9.563 | 5.131 | 4.433 | 19.649 | 46% |
| 16 | 9.865 | 5.478 | 4.387 | 19.249 | 44% |
| 17 | 10.899 | 6.148 | 4.751 | 22.575 | 44% |
| 18 | 10.751 | 5.657 | 5.094 | 25.947 | 47% |
| 19 | 12.242 | 5.506 | 6.737 | 45.382 | 55% |
| 20 | 8.288 | 5.253 | 3.036 | 9.215 | 37% |
| 21 | 7.503 | 4.991 | 2.511 | 6.307 | 33% |
| 22 | 7.314 | 4.710 | 2.603 | 6.778 | 36% |
| 23 | 7.554 | 6.791 | 0.763 | 0.582 | 10% |
| Rata-Rata | 16.649 | 7.746 | 8.903 | 151.410 | 47% |
| Jumlah | 382.926 | 178.155 | 204.770 | 3482.419 | - |

Sumber : Hasil Perhitungan

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.1 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{2.447}{10.720} \right| \times 100 \%$$

$$PE 1 = 23 \%$$

$$\blacksquare PE 2 = \left| \frac{4.740}{15.407} \right| \times 100 \%$$

$$PE 2 = 31 \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare PE 23 = \left| \frac{0.763}{7.554} \right| \times 100 \%$$

$$PE 23 = 10 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{23\% + 31\% \dots \dots 10\%}{23}$$

$$MAPE = 47 \%$$

- R^2

Didapatkan langsung dari software $R = 0.533$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.533} \\ &= 0.73 \end{aligned}$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah Tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R.

Tabel 4. 2 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|--------|-----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 10.720 | 11.764 | -1.044 | 1.089 | 10% |
| 2 | 15.407 | 7.696 | 7.711 | 59.457 | 50% |
| 3 | 21.262 | 9.470 | 11.792 | 139.047 | 55% |
| 4 | 35.495 | 9.399 | 26.096 | 681.022 | 74% |
| 5 | 44.300 | 8.033 | 36.267 | 1315.326 | 82% |
| 6 | 31.515 | 12.000 | 19.515 | 380.836 | 62% |
| 7 | 33.757 | 14.132 | 19.624 | 385.108 | 58% |
| 8 | 32.524 | 10.994 | 21.530 | 463.523 | 66% |
| 9 | 18.975 | 7.560 | 11.416 | 130.315 | 60% |
| 10 | 11.434 | 6.003 | 5.431 | 29.500 | 48% |
| 11 | 11.359 | 5.675 | 5.684 | 32.305 | 50% |
| 12 | 10.948 | 5.049 | 5.899 | 34.796 | 54% |
| 13 | 11.348 | 5.028 | 6.320 | 39.940 | 56% |
| 14 | 9.902 | 5.206 | 4.697 | 22.058 | 47% |
| 15 | 9.563 | 4.696 | 4.867 | 23.692 | 51% |
| 16 | 9.865 | 4.816 | 5.049 | 25.491 | 51% |
| 17 | 10.899 | 5.097 | 5.802 | 33.666 | 53% |

| | | | | | |
|------------------|----------------|----------------|----------------|-----------------|------------|
| 18 | 10.751 | 4.869 | 5.882 | 34.592 | 55% |
| 19 | 12.242 | 5.662 | 6.580 | 43.298 | 54% |
| 20 | 8.288 | 4.295 | 3.993 | 15.946 | 48% |
| 21 | 7.503 | 4.201 | 3.301 | 10.898 | 44% |
| 22 | 7.314 | 4.576 | 2.737 | 7.493 | 37% |
| 23 | 7.554 | 6.120 | 1.434 | 2.058 | 19% |
| Rata-Rata | 16.649 | 7.058 | 9.591 | 170.063 | 51% |
| Jumlah | 382.926 | 162.342 | 220.584 | 3911.455 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.2 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-1.044}{10.720} \right| \times 100 \%$

$$PE\ 1 = 10 \%$$

- $PE\ 2 = \left| \frac{7.711}{15.407} \right| \times 100 \%$

$$PE\ 2 = 50 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{1.434}{7.554} \right| \times 100 \%$

$$PE\ 23 = 19 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{10\% + 50\% \dots \dots 19\%}{23}$$

$$MAPE = 51 \%$$

- **R²**

Didapatkan langsung dari software R = 0.499

- **Korelasi**

$$\text{Korelasi} = \sqrt{R^2}$$

$$= \sqrt{0.499}$$

$$= 0.706$$

Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.3 :

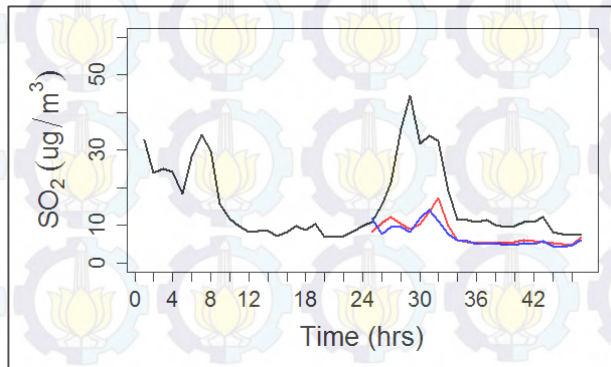
Tabel 4. 3 Performa Model ARIMA dan STL

| Data In Sample | | 4 Agustus 2014 - 18 Agustus 2014 | |
|--------------------------------|----------|----------------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 47% | 55% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.533 | 0.499 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.73 | 0.706 |

Berdasarkan Tabel 4.3 performa STL lebih bagus dibandingkan ARIMA karena MAPE pada Model STL lebih kecil daripada ARIMA yang artinya tingkat kesalahan ARIMA lebih besar daripada STL. Meskipun tingkat kesalahan model STL lebih kecil dibandingkan dengan ARIMA, tetapi nilai MAPE pada model tersebut masih dibawah standard yang ditetapkan. Penilaian performa model juga

dibuktikan dengan nilai R^2 dan korelasi. Model STL memiliki nilai R^2 dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.3 nilai R^2 STL sebesar 0.533 sedangkan ARIMA hanya 0.499

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.1



Gambar 4. 1 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter SO₂

Berdasarkan Gambar 4.1 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada grafik tersebut terlihat kedua model memiliki grafik yang hampir serupa, terbukti karena nilai R^2 kedua model yang tercantum pada Tabel 4.3 tidak terpaut jauh.

Model STL memiliki nilai R^2 dibawah standard yang ditetapkan karena masih dibawah 0.6. Terbukti pada gambar grafik tersebut, pola model STL tidak mampu mengikuti pola data verifikasi, Karena pada jam ke 26 hingga jam ke 32 (Nomor 1 hingga nomor 9 pada Tabel 4.1) model STL memiliki tingkat

kesalahan diatas 50%. Jadi dapat disimpulkan pada pengujian pertama kedua model tidak mampu untuk memprediksi zat pencemar SO_2 dikarenakan tingkat error pada hasil prediksi masih diatas 20% dan nilai R^2 masih dibawah 0.6

2. Data Tanggal 7 Agustus 2014 - 21 Agustus 2014

Pengujian model menggunakan data in sample pada tanggal 7 Agustus 2014 hingga tanggal 20 Agustus 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 21 Agustus 2014 Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (21 Agustus 2014).

Model terbaik adalah model dengan performa sesuai dengan standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model :

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.4 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | | |
|---|--------|--------|---------|-------------|-----|--|
| No | O | P | O - P | $(O - P)^2$ | PE | |
| 1 | 7.203 | 9.628 | - 2.425 | 5.879 | 34% | |
| 2 | 6.966 | 13.147 | - 6.182 | 38.216 | 89% | |
| 3 | 15.869 | 15.748 | 0.121 | 0.015 | 1% | |
| 4 | 32.731 | 14.444 | 18.287 | 334.396 | 56% | |
| 5 | 20.183 | 13.376 | 6.806 | 46.327 | 34% | |
| 6 | 10.670 | 15.142 | - 4.471 | 19.991 | 42% | |
| 7 | 16.487 | 19.672 | - 3.186 | 10.150 | 19% | |

| | | | | | |
|------------------|----------------|----------------|---------------|----------------|------------|
| 8 | 27.043 | 26.568 | 0.476 | 0.226 | 2% |
| 9 | 11.836 | 15.390 | - 3.554 | 12.632 | 30% |
| 10 | 8.649 | 8.424 | 0.225 | 0.051 | 3% |
| 11 | 8.709 | 7.190 | 1.518 | 2.306 | 17% |
| 12 | 9.444 | 6.778 | 2.667 | 7.111 | 28% |
| 13 | 9.718 | 6.720 | 2.997 | 8.983 | 31% |
| 14 | 9.982 | 6.553 | 3.430 | 11.763 | 34% |
| 15 | 10.079 | 6.376 | 3.703 | 13.709 | 37% |
| 16 | 10.003 | 6.820 | 3.183 | 10.131 | 32% |
| 17 | 12.872 | 7.672 | 5.200 | 27.041 | 40% |
| 18 | 12.135 | 7.100 | 5.034 | 25.345 | 41% |
| 19 | 7.565 | 7.113 | 0.451 | 0.204 | 6% |
| 20 | 6.703 | 6.083 | 0.620 | 0.385 | 9% |
| 21 | 6.512 | 5.695 | 0.817 | 0.667 | 13% |
| 22 | 7.286 | 5.428 | 1.858 | 3.450 | 25% |
| 23 | 6.798 | 7.598 | - 0.800 | 0.640 | 12% |
| Rata-Rata | 11.976 | 10.377 | 1.599 | 25.201 | 28% |
| Jumlah | 275.443 | 238.668 | 36.776 | 579.617 | - |

Sumber : Hasil Perhitungan

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.4 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-2.425}{7.203} \right| \times 100\ %$

$$PE\ 1 = 34\ %$$

- $PE\ 2 = \left| \frac{-6.182}{6.996} \right| \times 100\ %$

$$PE\ 2 = 89\ %$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-0.800}{6.798} \right| \times 100\ %$

$$PE\ 23 = 12\ %$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{34\% + 89\% \dots \dots 12\%}{23}$$

$$MAPE = 28\ %$$

- **R^2**

Didapatkan langsung dari software $R = 0.483$

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.483}$$

$$= 0.695$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai ME, RMSE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Plot ACF dan PACF dapat dilihat di lihat pada gambar berikut ini

Berdasarkan gambar di lampiran dapat dilihat data sudah stasioner. Hal itu ditunjukkan dengan data menurun secara cepat (Dies down) dan membentuk pola sinusoidal. Pada pengujian ini digunakan perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman) Transformasi Box-Cox dilakukan untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima.

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.5 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|---------|-----------------------|------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 7.203 | 10.986 | -3.783 | 14.312 | 53% |
| 2 | 6.966 | 13.326 | -6.360 | 40.451 | 91% |
| 3 | 15.869 | 14.356 | 1.514 | 2.291 | 10% |
| 4 | 32.731 | 11.417 | 21.314 | 454.278 | 65% |
| 5 | 20.183 | 10.509 | 9.673 | 93.571 | 48% |
| 6 | 10.670 | 17.292 | -6.621 | 43.838 | 62% |
| 7 | 16.487 | 19.432 | -2.946 | 8.679 | 18% |
| 8 | 27.043 | 29.924 | -2.880 | 8.295 | 11% |
| 9 | 11.836 | 25.104 | -13.268 | 176.050 | 112% |

| | | | | | |
|------------------|----------------|----------------|----------------|-----------------|------------|
| 10 | 8.649 | 16.459 | -7.810 | 60.998 | 90% |
| 11 | 8.709 | 13.525 | -4.816 | 23.192 | 55% |
| 12 | 9.444 | 12.762 | -3.317 | 11.004 | 35% |
| 13 | 9.718 | 12.466 | -2.749 | 7.555 | 28% |
| 14 | 9.982 | 12.458 | -2.476 | 6.129 | 25% |
| 15 | 10.079 | 12.373 | -2.294 | 5.264 | 23% |
| 16 | 10.003 | 12.454 | -2.451 | 6.008 | 25% |
| 17 | 12.872 | 12.382 | 0.491 | 0.241 | 4% |
| 18 | 12.135 | 12.542 | -0.407 | 0.166 | 3% |
| 19 | 7.565 | 12.767 | -5.203 | 27.066 | 69% |
| 20 | 6.703 | 11.552 | -4.848 | 23.506 | 72% |
| 21 | 6.512 | 11.350 | -4.838 | 23.411 | 74% |
| 22 | 7.286 | 11.401 | - 4.115 | 16.936 | 56% |
| 23 | 6.798 | 12.153 | -5.355 | 28.675 | 79% |
| Rata-Rata | 11.976 | 14.304 | -2.328 | 47.040 | 48% |
| Jumlah | 275.443 | 328.990 | -53.547 | 1081.915 | - |

Sumber : Hasil Perhitungan

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.4 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{-3.783}{7.203} \right| \times 100 \%$$

$$PE\ 1 = 53\ \%$$

$$\blacksquare\ PE\ 2 = \left| \frac{-6.360}{6.996} \right| \times 100\ \%$$

$$PE\ 2 = 91\ \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare\ PE\ 23 = \left| \frac{-5.355}{6.798} \right| \times 100\ \%$$

$$PE\ 23 = 79\ \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{53\% + 91\% \dots \dots 79\%}{23}$$

$$MAPE = 48\ \%$$

- R^2

Didapatkan langsung dari software $R = 0.698$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.698} \\ &= 0.835 \end{aligned}$$

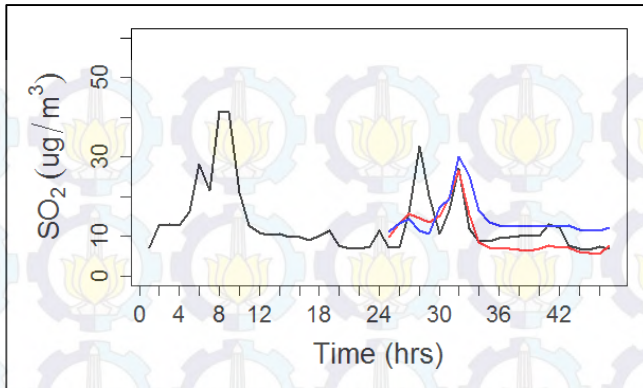
Performa kedua model pada pengujian kedua dapat dilihat melalui Tabel 4.6 :

Tabel 4.6 Performa Model ARIMA dan STL

| Data In Sample | | 7 Agustus 2014 - 21 Agustus 2014 | |
|--------------------------------|----------|----------------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 28% | 48% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.483 | 0.138 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.695 | 0.371 |

Berdasarkan Tabel 4.6 performa STL lebih bagus dibandingkan ARIMA karena tingkat kesalahan (MAPE) pada model STL lebih sedikit daripada model ARIMA akan tetapi, meskipun nilai MAPE pada model STL lebih baik daripada ARIMA Model STL tidak cukup baik dalam melakukan prediksi pada pencemar SO₂ karena memiliki nilai MAPE diatas 20%.Penilaian performa model juga dibuktikan dengan nilai R² dan korelasi. Model STL memiliki nilai R² dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.6 nilai R² STL sebesar 0.483 sedangkan ARIMA hanya 0.138

Nilai R² dan korelasi menunjukan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.2



Gambar 4. 2 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter SO₂

Berdasarkan Gambar 4.2 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli).

Model STL dan Model ARIMA memiliki nilai R^2 dibawah standard yang ditetapkan karena masih dibawah 0.6. Terbukti pada gambar grafik tersebut terlihat kedua model tidak mampu mengikuti pola data verifikasi pada jam ke-28 (atau nomor 4 pada Tabel 4.4 dan nomor 4 pada Tabel 4.5) model STL dan ARIMA memprediksi jauh lebih kecil dari data sebenarnya, sehingga mengakibatkan pada waktu tersebut kedua model memiliki nilai error diatas 50%. Jadi dapat disimpulkan pada pengujian kedua, kedua model tidak mampu untuk memprediksi zat pencemar SO₂ dikarenakan tingkat error pada hasil prediksi masih diatas 20% dan nilai R^2 masih dibawah 0.6

3. Data Tanggal 10 Agustus 2014 - 24 Agustus 2014

Pengujian model menggunakan data in sample pada tanggal 10 Agustus 2014 hingga tanggal 23 Agustus 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 24 Agustus 2014. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (24 Agustus 2014). Model terbaik adalah model dengan performa sesuai dengan standard kriteria yang telah ditentukan pada bab sebelumnya.

Berikut adalah analisis dan hasil prediksi masing masing model:

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.7 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|---------|----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 12.989 | 12.847 | 0.142 | 0.020 | 1% |
| 2 | 14.307 | 15.329 | - 1.022 | 1.044 | 7% |
| 3 | 14.460 | 18.788 | - 4.328 | 18.735 | 30% |
| 4 | 15.810 | 20.960 | - 5.149 | 26.515 | 33% |
| 5 | 14.378 | 19.061 | - 4.683 | 21.928 | 33% |
| 6 | 11.837 | 19.635 | - 7.799 | 60.820 | 66% |
| 7 | 17.022 | 24.859 | - 7.837 | 61.416 | 46% |
| 8 | 58.653 | 32.506 | 26.148 | 683.710 | 45% |
| 9 | 25.235 | 21.060 | 4.175 | 17.435 | 17% |
| 10 | 20.129 | 12.936 | 7.194 | 51.751 | 36% |
| 11 | 10.580 | 11.691 | - 1.112 | 1.236 | 11% |
| 12 | 7.681 | 11.262 | - 3.581 | 12.822 | 47% |
| 13 | 7.845 | 11.289 | - 3.443 | 11.856 | 44% |
| 14 | 7.613 | 11.165 | - 3.552 | 12.616 | 47% |
| 15 | 7.976 | 10.977 | - 3.001 | 9.009 | 38% |

| | | | | | |
|------------------|----------------|----------------|-----------------|------------------|------------|
| 16 | 9.473 | 11.298 | - 1.825 | 3.330 | 19% |
| 17 | 9.978 | 12.271 | - 2.293 | 5.258 | 23% |
| 18 | 8.907 | 11.676 | - 2.769 | 7.667 | 31% |
| 19 | 8.246 | 11.232 | - 2.987 | 8.921 | 36% |
| 20 | 8.718 | 9.961 | - 1.243 | 1.545 | 14% |
| 21 | 6.770 | 9.643 | - 2.873 | 8.256 | 42% |
| 22 | 8.459 | 9.455 | - 0.996 | 0.993 | 12% |
| 23 | 7.220 | 11.206 | - 3.987 | 15.893 | 55% |
| Rata-Rata | 13.665 | 14.831 | - 1.166 | 45.338 | 32% |
| Jumlah | 314.287 | 341.107 | - 26.821 | 1,042.775 | - |

Sumber : Hasil Perhitungan

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah `auto.arima`, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.8 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|---------|---------|-----------|-----------------------|------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 12.989 | 15.429 | - 2.440 | 5.952 | 19% |
| 2 | 14.307 | 17.683 | - 3.376 | 11.400 | 24% |
| 3 | 14.460 | 20.271 | - 5.811 | 33.771 | 40% |
| 4 | 15.810 | 22.200 | - 6.389 | 40.820 | 40% |
| 5 | 14.378 | 20.944 | - 6.566 | 43.109 | 46% |
| 6 | 11.837 | 21.327 | - 9.490 | 90.063 | 80% |
| 7 | 17.022 | 24.735 | - 7.713 | 59.489 | 45% |
| 8 | 58.653 | 29.515 | 29.138 | 849.029 | 50% |
| 9 | 25.235 | 22.520 | 2.715 | 7.372 | 11% |
| 10 | 20.129 | 17.124 | 3.005 | 9.030 | 15% |
| 11 | 10.580 | 16.285 | - 5.705 | 32.547 | 54% |
| 12 | 7.681 | 15.981 | - 8.299 | 68.879 | 108% |
| 13 | 7.845 | 15.969 | - 8.124 | 65.992 | 104% |
| 14 | 7.613 | 15.906 | - 8.293 | 68.776 | 109% |
| 15 | 7.976 | 15.805 | - 7.829 | 61.300 | 98% |
| 16 | 9.473 | 15.955 | - 6.481 | 42.009 | 68% |
| 17 | 9.978 | 16.479 | - 6.501 | 42.264 | 65% |
| 18 | 8.907 | 16.182 | - 7.275 | 52.921 | 82% |
| 19 | 8.246 | 15.869 | - 7.623 | 58.114 | 92% |
| 20 | 8.718 | 15.057 | - 6.339 | 40.183 | 73% |
| 21 | 6.770 | 14.923 | - 8.152 | 66.463 | 120% |
| 22 | 8.459 | 14.783 | - 6.324 | 39.992 | 75% |
| 23 | 7.220 | 15.789 | - 8.570 | 73.442 | 119% |
| Rata-Rata | 13.665 | 18.119 | - 4.454 | 80.997 | 67% |
| Jumlah | 314.287 | 416.730 | - 102.443 | 1,862.920 | - |

Sumber : Hasil Perhitungan

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.8 dapat diketahui

• **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE\ 1 = \left| \frac{-2.440}{12.989} \right| \times 100 \%$$

$$PE\ 1 = 19 \%$$

$$\blacksquare PE\ 2 = \left| \frac{-3.376}{14.307} \right| \times 100 \%$$

$$PE\ 2 = 24 \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare PE\ 23 = \left| \frac{-0.857}{7.220} \right| \times 100 \%$$

$$PE\ 23 = 119 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{34\% + 89\% \dots \dots 119\%}{23}$$

$$MAPE = 67 \%$$

- **R²**

Didapatkan langsung dari software $R = 0.688$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.688} \\ &= 0.829 \end{aligned}$$

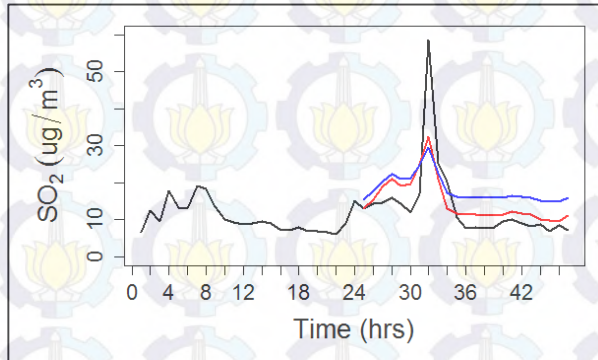
Performa kedua model pada pengujian Ketiga dapat dilihat melalui tabel berikut ini :

Tabel 4.9 Performa Model ARIMA dan STL

| Data In Sample | | 10 Agustus 2014 - 24 Agustus 2014 | |
|--------------------------------|----------|-----------------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 32% | 67% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.698 | 0.688 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.835 | 0.829 |

Berdasarkan tabel 4.9 performa STL lebih bagus dibandingkan ARIMA karena tingkat kesalahan (MAPE) pada model STL lebih sedikit daripada model ARIMA akan tetapi, meskipun nilai MAPE pada model STL lebih baik daripada ARIMA Model STL tidak cukup baik dalam melakukan prediksi pada pencemar SO₂ karena memiliki nilai MAPE diatas 20%. Penilaian performa model juga dibuktikan dengan nilai R² dan korelasi. Model STL memiliki nilai R² dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.6 nilai R² STL sebesar 0.698 sedangkan ARIMA sebesar 0.688.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.3



Gambar 4. 3 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter SO_2

Berdasarkan Gambar 4.3 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli).

Model STL dan Model ARIMA memiliki nilai R^2 sesuai dengan standard yang ditetapkan karena lebih dari 0.6. Terbukti pada gambar grafik tersebut terlihat kedua model mampu mengikuti pola data verifikasi, meskipun terdapat kesalahan pada beberapa jam. Kesalahan paling signifikan terlihat pada jam ke-32 (atau nomor 8 pada Tabel 4.7 dan Tabel 4.8) model STL dan ARIMA memprediksi jauh lebih kecil dari data sebenarnya, sehingga mengakibatkan pada waktu tersebut kedua model memiliki nilai error diatas 40%. Kesalahan terbesar dilakukan oleh model ARIMA karena memiliki kesalahan diatas 100%, hal ini terjadi karena hasil prediksi dua kali lebih besar dari data asli,

sehingga tingkat error juga dua kali lebih besar. Terbukti pada jam ke 47 atau nomor 23 pada Tabel 4.8 hasil prediksi ARIMA sebesar 15.789 sedangkan seharusnya hanya 7.220. hal ini menimbulkan tingkat kesalahan prediksi sebesar 119%

Jadi dapat disimpulkan dari ketiga pengujian, kedua model tidak mampu untuk memprediksi zat pencemar SO_2 dikarenakan tingkat error pada hasil prediksi masih diatas 20% dan nilai R^2 masih tidak konsisten diatas 0.6

4.1.2. CO

1. Data Tanggal 12 Juli 2014 - 26 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 12 Juli 2014 hingga tanggal 25 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 26 Juli 2014. Karakteristik zat pencemar CO pada tanggal 12 Juli 2014 hingga tanggal 25 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (26 Juli 2014).

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model :

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.10 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|-------|-------|--------|-----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 0.436 | 0.456 | -0.020 | 0.0004 | 5% |
| 2 | 0.331 | 0.404 | -0.073 | 0.0054 | 22% |
| 3 | 0.255 | 0.367 | -0.112 | 0.0126 | 44% |
| 4 | 0.243 | 0.354 | -0.111 | 0.0124 | 46% |

| | | | | | |
|------------------|--------------|--------------|----------------|---------------|------------|
| 5 | 0.257 | 0.297 | -0.040 | 0.0016 | 16% |
| 6 | 0.346 | 0.390 | -0.044 | 0.0020 | 13% |
| 7 | 0.548 | 0.532 | 0.016 | 0.0003 | 3% |
| 8 | 0.611 | 0.622 | -0.011 | 0.0001 | 2% |
| 9 | 0.556 | 0.808 | -0.253 | 0.0639 | 46% |
| 10 | 0.450 | 0.674 | -0.223 | 0.0498 | 50% |
| 11 | 0.352 | 0.499 | -0.147 | 0.0216 | 42% |
| 12 | 0.315 | 0.382 | -0.068 | 0.0046 | 22% |
| 13 | 0.269 | 0.292 | -0.023 | 0.0005 | 9% |
| 14 | 0.238 | 0.240 | -0.002 | 0.0000 | 1% |
| 15 | 0.236 | 0.211 | 0.025 | 0.0006 | 11% |
| 16 | 0.282 | 0.260 | 0.022 | 0.0005 | 8% |
| 17 | 0.345 | 0.252 | 0.093 | 0.0087 | 27% |
| 18 | 0.407 | 0.477 | -0.070 | 0.0050 | 17% |
| 19 | 0.378 | 0.370 | 0.008 | 0.0001 | 2% |
| 20 | 0.393 | 0.433 | -0.040 | 0.0016 | 10% |
| 21 | 0.445 | 0.493 | -0.049 | 0.0024 | 11% |
| 22 | 0.438 | 0.330 | 0.108 | 0.0116 | 25% |
| 23 | 0.443 | 0.400 | 0.043 | 0.0018 | 10% |
| Rata-Rata | 0.373 | 0.400 | - 0.042 | 0.0090 | 19% |
| Jumlah | 8.571 | 9.544 | - 0.973 | 0.207 | - |

Sumber : Hasil Perhitungan

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.10 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-0.020}{0.436} \right| \times 100 \%$

$$PE\ 1 = 5 \%$$

- $PE\ 2 = \left| \frac{-0.073}{0.331} \right| \times 100 \%$

$$PE\ 2 = 22 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{0.043}{0.443} \right| \times 100 \%$

$$PE\ 23 = 10 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{5\% + 22\% \dots \dots 10\%}{23}$$

$$MAPE = 19 \%$$

- **R²**

Didapatkan langsung dari software R = 0.788

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.788}$$

$$= 0.888$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.11 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|-------|-------|--------|-----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 0.436 | 0.455 | -0.019 | 0.0004 | 4% |
| 2 | 0.331 | 0.422 | -0.091 | 0.0083 | 28% |
| 3 | 0.255 | 0.393 | -0.139 | 0.0192 | 54% |
| 4 | 0.243 | 0.395 | -0.152 | 0.0232 | 63% |
| 5 | 0.257 | 0.410 | -0.153 | 0.0234 | 60% |
| 6 | 0.346 | 0.458 | -0.112 | 0.0126 | 32% |
| 7 | 0.548 | 0.597 | -0.050 | 0.0025 | 9% |
| 8 | 0.611 | 0.818 | -0.208 | 0.0431 | 34% |
| 9 | 0.556 | 0.940 | -0.384 | 0.1478 | 69% |
| 10 | 0.450 | 0.798 | -0.347 | 0.1205 | 77% |
| 11 | 0.352 | 0.633 | -0.280 | 0.0786 | 80% |
| 12 | 0.315 | 0.503 | -0.188 | 0.0354 | 60% |
| 13 | 0.269 | 0.411 | -0.142 | 0.0202 | 53% |

| | | | | | |
|-----------|-------|--------|---------|--------|-----|
| 14 | 0.238 | 0.363 | -0.125 | 0.0155 | 52% |
| 15 | 0.236 | 0.355 | -0.119 | 0.0141 | 50% |
| 16 | 0.282 | 0.369 | -0.087 | 0.0076 | 31% |
| 17 | 0.345 | 0.420 | -0.075 | 0.0057 | 22% |
| 18 | 0.407 | 0.508 | -0.102 | 0.0103 | 25% |
| 19 | 0.378 | 0.546 | -0.168 | 0.0283 | 45% |
| 20 | 0.393 | 0.583 | -0.190 | 0.0362 | 48% |
| 21 | 0.445 | 0.635 | -0.190 | 0.0361 | 43% |
| 22 | 0.438 | 0.676 | -0.239 | 0.0570 | 55% |
| 23 | 0.443 | 0.675 | -0.232 | 0.0540 | 52% |
| Rata-Rata | 0.373 | 0.538 | - 0.165 | 0.035 | 45% |
| Jumlah | 8.571 | 12.363 | - 3.792 | 0.800 | - |

Sumber : Hasil Penelitian

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.11 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE_1 = \left| \frac{-0.019}{0.436} \right| \times 100 \%$

$$PE_1 = 4 \%$$

- $PE_2 = \left| \frac{-0.091}{0.331} \right| \times 100 \%$

$$PE\ 2 = 28\ \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare\ PE\ 23 = \left| \frac{-0.232}{0.443} \right| \times 100\ \%$$

$$PE\ 23 = 52\ \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{4\% + 28\% \dots \dots 52\%}{23}$$

$$MAPE = 45\ \%$$

- R^2

Didapatkan langsung dari software $R = 0.736$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.736} \\ &= 0.857 \end{aligned}$$

Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.12

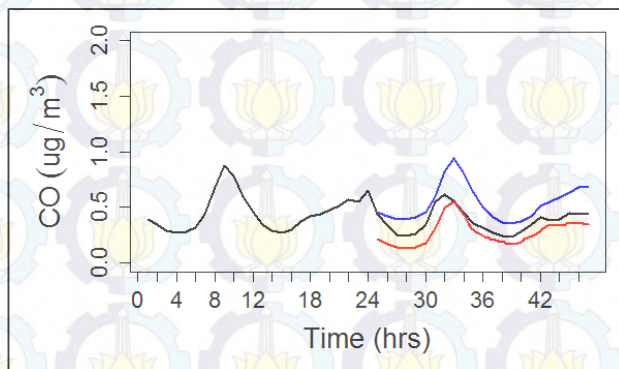
Tabel 4.12 Performa Model ARIMA dan STL

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 19% | 45% |
| R^2 | STANDARD | > 0.6 | |
| | HASIL | 0.788 | 0.736 |

| Correlation | STANDARD | (+ 0.774) | |
|-------------|----------|------------|-------|
| | HASIL | 0.888 | 0.858 |

Berdasarkan Tabel 4.12 performa STL lebih bagus dibandingkan ARIMA karena MAPE pada Model STL lebih kecil daripada ARIMA dan sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model STL memiliki nilai R^2 dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.12 nilai R^2 STL sebesar 0.788 sedangkan ARIMA hanya 0.736

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.4



Gambar 4. 4 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter CO

Berdasarkan Gambar 4.4 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada grafik tersebut terlihat kedua model memiliki grafik yang hampir serupa, terbukti karena nilai R^2 kedua model yang tercantum pada Tabel 4.12 tidak terpaut jauh.

Model STL memiliki nilai R^2 sesuai dengan standard yang telah ditetapkan karena diatas 0.6. Terbukti pada gambar grafik tersebut, pola model STL mampu mengikuti pola data verifikasi,

Jadi dapat disimpulkan pada pengujian pertama kedua model mampu untuk memprediksi zat pencemar CO dikarenakan tingkat error pada hasil prediksi dibawah 20% dan nilai R^2 diatas 0.6.

2. Data Tanggal 15 Juli 2014 - 29 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 15 Juli 2014 hingga tanggal 28 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 29 Juli 2014. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (29 Juli 2014). Model terbaik adalah model dengan performa diatas standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model :

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.13 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|-------|-------|---------|-----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 0.255 | 0.212 | 0.042 | 0.0018 | 17% |
| 2 | 0.269 | 0.172 | 0.098 | 0.0095 | 36% |
| 3 | 0.304 | 0.300 | 0.004 | 0.00001 | 1% |
| 4 | 0.311 | 0.290 | 0.021 | 0.0005 | 7% |
| 5 | 0.309 | 0.330 | - 0.021 | 0.0005 | 7% |
| 6 | 0.344 | 0.320 | 0.024 | 0.0006 | 7% |
| 7 | 0.422 | 0.390 | 0.032 | 0.0010 | 8% |

| | | | | | |
|------------------|--------------|--------------|--------------|--------------|------------|
| 8 | 0.445 | 0.446 | - 0.001 | 0.0000 | 0.2% |
| 9 | 0.452 | 0.455 | - 0.003 | 0.00001 | 1% |
| 10 | 0.359 | 0.330 | 0.029 | 0.0008 | 8% |
| 11 | 0.278 | 0.279 | - 0.001 | 0.00001 | 1% |
| 12 | 0.230 | 0.233 | - 0.003 | 0.00001 | 1% |
| 13 | 0.208 | 0.098 | 0.110 | 0.0120 | 53% |
| 14 | 0.185 | 0.056 | 0.129 | 0.0166 | 70% |
| 15 | 0.180 | 0.178 | 0.002 | 0.00001 | 1% |
| 16 | 0.183 | 0.177 | 0.006 | 0.00001 | 3% |
| 17 | 0.197 | 0.190 | 0.007 | 0.00001 | 4% |
| 18 | 0.222 | 0.205 | 0.017 | 0.0003 | 8% |
| 19 | 0.247 | 0.249 | - 0.002 | 0.00001 | 1% |
| 20 | 0.287 | 0.286 | 0.002 | 0.00001 | 1% |
| 21 | 0.315 | 0.320 | - 0.005 | 0.00001 | 2% |
| 22 | 0.334 | 0.337 | - 0.003 | 0.00001 | 1% |
| 23 | 0.326 | 0.329 | - 0.003 | 0.00001 | 1% |
| Rata-Rata | 0.290 | 0.269 | 0.021 | 0.002 | 10% |
| Jumlah | 6.662 | 6.182 | 0.480 | 0.044 | - |

Sumber : Hasil Perhitungan

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.13 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{0.042}{0.255} \right| \times 100\ %$

$$PE\ 1 = 17\ %$$

- $PE\ 2 = \left| \frac{0.098}{0.269} \right| \times 100\ %$

$$PE\ 2 = 26\ %$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-0.003}{0.326} \right| \times 100\ %$

$$PE\ 23 = 1\ %$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{17\% + 26\% \dots \dots 1\%}{23}$$

$$MAPE = 10\ %$$

- R^2

Didapatkan langsung dari software $R = 0.853$

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.853}$$

$$= 0.923$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat

stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.14 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|-------|-------|---------|-----------------------|-----|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 0.255 | 0.212 | 0.042 | 0.002 | 17% |
| 2 | 0.269 | 0.172 | 0.098 | 0.010 | 36% |
| 3 | 0.304 | 0.136 | 0.168 | 0.028 | 55% |
| 4 | 0.311 | 0.129 | 0.182 | 0.033 | 58% |
| 5 | 0.309 | 0.139 | 0.170 | 0.029 | 55% |
| 6 | 0.344 | 0.177 | 0.168 | 0.028 | 49% |
| 7 | 0.422 | 0.307 | 0.115 | 0.013 | 27% |
| 8 | 0.445 | 0.498 | - 0.053 | 0.003 | 12% |
| 9 | 0.452 | 0.551 | - 0.099 | 0.010 | 22% |
| 10 | 0.359 | 0.426 | - 0.067 | 0.004 | 19% |
| 11 | 0.278 | 0.279 | - 0.001 | 0.000 | 1% |
| 12 | 0.230 | 0.177 | 0.053 | 0.003 | 23% |
| 13 | 0.208 | 0.098 | 0.110 | 0.012 | 53% |
| 14 | 0.185 | 0.056 | 0.129 | 0.017 | 70% |
| 15 | 0.180 | 0.046 | 0.134 | 0.018 | 75% |
| 16 | 0.183 | 0.061 | 0.121 | 0.015 | 66% |

| | | | | | |
|------------------|--------------|--------------|--------------|--------------|------------|
| 17 | 0.197 | 0.110 | 0.087 | 0.008 | 44% |
| 18 | 0.222 | 0.205 | 0.017 | 0.000 | 8% |
| 19 | 0.247 | 0.249 | - 0.002 | 0.000 | 1% |
| 20 | 0.287 | 0.286 | 0.002 | 0.000 | 1% |
| 21 | 0.315 | 0.344 | - 0.030 | 0.001 | 9% |
| 22 | 0.334 | 0.399 | - 0.065 | 0.004 | 19% |
| 23 | 0.326 | 0.426 | - 0.100 | 0.010 | 31% |
| Rata-Rata | 0.290 | 0.238 | 0.051 | 0.011 | 33% |
| Jumlah | 6.662 | 5.483 | 1.179 | 0.247 | - |

Sumber : Hasil Penelitian

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.14 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{0.042}{0.255} \right| \times 100 \%$

$$PE\ 1 = 17 \%$$

- $PE\ 2 = \left| \frac{0.098}{0.269} \right| \times 100 \%$

$$PE\ 2 = 36 \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare PE_{23} = \left| \frac{-0.1}{0.326} \right| \times 100 \%$$

$$PE_{23} = 31 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{17\% + 36\% \dots \dots 31\%}{23}$$

$$MAPE = 33 \%$$

- R^2

Didapatkan langsung dari software $R = 0.628$

- Korelasi

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.628} \\ &= 0.792 \end{aligned}$$

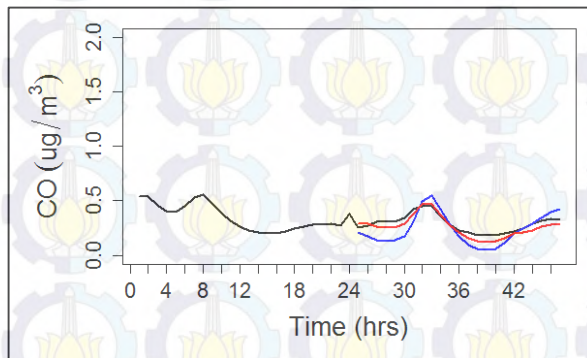
Performa kedua model pada pengujian pertama dapat dilihat melalui tabel berikut ini :

Tabel 4.15 Performa Model ARIMA dan STL

| Data In Sample | | 15 Juli 2014 - 29 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 10% | 33% |
| R^2 | STANDARD | > 0.6 | |
| | HASIL | 0.853 | 0.628 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.924 | 0.792 |

Berdasarkan Tabel 4.15 performa STL lebih bagus dibandingkan ARIMA karena MAPE pada Model STL lebih kecil daripada ARIMA dan sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model STL memiliki nilai R^2 dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.12 nilai R^2 STL sebesar 0.853 sedangkan ARIMA hanya 0.628

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.5



Gambar 4. 5 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter CO

Berdasarkan Gambar 4.5 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada grafik tersebut terlihat model STL memiliki grafik yang hampir serupa dengan data asli, terbukti dengan nilai R^2 model yang tercantum pada Tabel 4.12 sebesar 0.853 dan tingkat error hanya 10%

Model STL memiliki nilai R^2 sesuai dengan standard yang telah ditetapkan karena diatas 0.6. Terbukti pada gambar grafik tersebut, pola model STL mampu mengikuti pola data verifikasi,

Jadi dapat disimpulkan pada pengujian pertama kedua model mampu untuk memprediksi zat pencemar CO dikarenakan tingkat error pada hasil prediksi dibawah 20% dan nilai R^2 diatas 0.6.

3. Data Tanggal 12 Juli 2014 - 26 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 18 Juli 2014 hingga tanggal 31 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 1 Agustus 2014. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (1 Agustus 2014). Model terbaik adalah model dengan performa diatas standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model :

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4. 16 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|-------|-------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 0.202 | 0.200 | 0.002 | 0.000 | 1% |
| 2 | 0.172 | 0.179 | - 0.007 | 0.000 | 4% |
| 3 | 0.169 | 0.164 | 0.005 | 0.000 | 3% |
| 4 | 0.199 | 0.170 | 0.029 | 0.001 | 15% |
| 5 | 0.223 | 0.193 | 0.031 | 0.001 | 14% |
| 6 | 0.278 | 0.244 | 0.034 | 0.001 | 12% |
| 7 | 0.370 | 0.370 | 0.000 | 0.000 | 0.02% |
| 8 | 0.471 | 0.510 | - 0.039 | 0.002 | 8% |
| 9 | 0.509 | 0.527 | - 0.018 | 0.000 | 4% |
| 10 | 0.407 | 0.391 | 0.017 | 0.000 | 4% |

| | | | | | |
|------------------|--------------|--------------|--------------|--------------|------------|
| 11 | 0.313 | 0.267 | 0.047 | 0.002 | 15% |
| 12 | 0.267 | 0.176 | 0.090 | 0.008 | 34% |
| 13 | 0.236 | 0.115 | 0.120 | 0.014 | 51% |
| 14 | 0.202 | 0.084 | 0.118 | 0.014 | 58% |
| 15 | 0.200 | 0.086 | 0.114 | 0.013 | 57% |
| 16 | 0.191 | 0.105 | 0.086 | 0.007 | 45% |
| 17 | 0.216 | 0.147 | 0.069 | 0.005 | 32% |
| 18 | 0.267 | 0.231 | 0.036 | 0.001 | 14% |
| 19 | 0.277 | 0.266 | 0.011 | 0.000 | 4% |
| 20 | 0.303 | 0.297 | 0.005 | 0.000 | 2% |
| 21 | 0.304 | 0.342 | - 0.038 | 0.001 | 13% |
| 22 | 0.312 | 0.384 | - 0.073 | 0.005 | 23% |
| 23 | 0.325 | 0.409 | - 0.084 | 0.007 | 26% |
| Rata-Rata | 0.279 | 0.255 | 0.024 | 0.004 | 16% |
| Jumlah | 6.413 | 5.859 | 0.554 | 0.084 | - |

Sumber :Hasil Peneitian

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.16 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{0.002}{0.202} \right| \times 100 \%$

$$PE\ 1 = 1\%$$

$$PE\ 2 = \left| \frac{-0.007}{0.172} \right| \times 100\%$$

$$PE\ 2 = 4\%$$

Dilanjutkan sampai PE ke 23

$$PE\ 23 = \left| \frac{-0.084}{0.325} \right| \times 100\%$$

$$PE\ 23 = 26\%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{1\% + 4\% \dots \dots 26\%}{23}$$

$$MAPE = 16\%$$

- R^2

Didapatkan langsung dari software $R = 0.628$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.628} \\ &= 0.792 \end{aligned}$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai ME, RMSE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan

menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4. 17 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|-------|-------|---------|-----------------------|------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 0.202 | 0.262 | - 0.060 | 0.004 | 30% |
| 2 | 0.172 | 0.256 | - 0.084 | 0.007 | 49% |
| 3 | 0.169 | 0.254 | - 0.085 | 0.007 | 50% |
| 4 | 0.199 | 0.257 | - 0.057 | 0.003 | 29% |
| 5 | 0.223 | 0.270 | - 0.047 | 0.002 | 21% |
| 6 | 0.278 | 0.304 | - 0.026 | 0.001 | 9% |
| 7 | 0.370 | 0.371 | - 0.001 | 0.000 | 0.3% |
| 8 | 0.471 | 0.418 | 0.053 | 0.003 | 11% |
| 9 | 0.509 | 0.401 | 0.107 | 0.012 | 21% |
| 10 | 0.407 | 0.326 | 0.081 | 0.007 | 20% |
| 11 | 0.313 | 0.261 | 0.052 | 0.003 | 17% |
| 12 | 0.267 | 0.220 | 0.047 | 0.002 | 18% |
| 13 | 0.236 | 0.189 | 0.046 | 0.002 | 20% |
| 14 | 0.202 | 0.170 | 0.033 | 0.001 | 16% |
| 15 | 0.200 | 0.162 | 0.038 | 0.001 | 19% |
| 16 | 0.191 | 0.166 | 0.025 | 0.001 | 13% |
| 17 | 0.216 | 0.177 | 0.039 | 0.002 | 18% |
| 18 | 0.267 | 0.196 | 0.071 | 0.005 | 26% |
| 19 | 0.277 | 0.209 | 0.068 | 0.005 | 25% |
| 20 | 0.303 | 0.229 | 0.074 | 0.005 | 24% |

| | | | | | |
|------------------|--------------|--------------|--------------|--------------|------------|
| 21 | 0.304 | 0.250 | 0.054 | 0.003 | 18% |
| 22 | 0.312 | 0.261 | 0.051 | 0.003 | 16% |
| 23 | 0.325 | 0.259 | 0.066 | 0.004 | 20% |
| Rata-Rata | 0.279 | 0.255 | 0.024 | 0.004 | 21% |
| Jumlah | 6.413 | 5.867 | 0.546 | 0.082 | - |

Sumber : Hasil Penelitian

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.17 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-0.060}{0.202} \right| \times 100 \%$

$$PE\ 1 = 30 \%$$

- $PE\ 2 = \left| \frac{-0.084}{0.172} \right| \times 100 \%$

$$PE\ 2 = 49 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{0.066}{0.325} \right| \times 100 \%$

$$PE\ 23 = 20 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{30\% + 49\% \dots \dots 20\%}{23}$$

$$MAPE = 21 \%$$

- **R²**

Didapatkan langsung dari software $R = 0.628$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.628} \\ &= 0.792 \\ &= 0.792 \end{aligned}$$

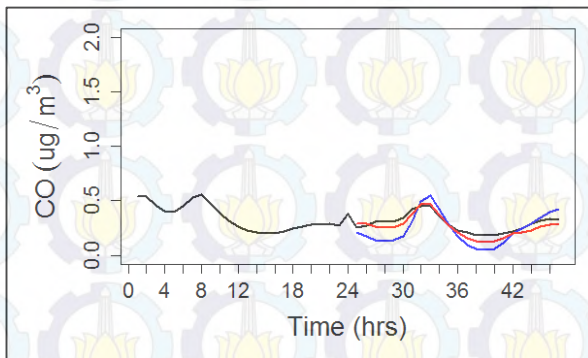
Performa kedua model pada pengujian pertama dapat dilihat melalui tabel berikut ini :

Tabel 4.18 Performa Model ARIMA dan STL

| Data In Sample | | 15 Juli 2014 - 29 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 16% | 21% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.853 | 0.628 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.924 | 0.792 |

Berdasarkan Tabel 4.18 performa STL lebih bagus dibandingkan ARIMA karena MAPE pada Model STL lebih kecil daripada ARIMA dan sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model STL memiliki nilai R^2 dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.18 nilai R^2 STL sebesar 0.853 sedangkan ARIMA hanya 0.628

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.5



Gambar 4. 6 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter CO

Berdasarkan Gambar 4.18 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada grafik tersebut terlihat model STL memiliki grafik yang hampir serupa dengan data asli, terbukti dengan nilai R^2 model yang tercantum pada Tabel 4.12 sebesar 0.853 dan tingkat error hanya 16%

Model STL memiliki nilai R^2 sesuai dengan standard yang telah ditetapkan karena diatas 0.6. Terbukti pada gambar grafik tersebut, pola model STL mampu mengikuti pola data verifikasi,

Jadi dapat disimpulkan pada pengujian pertama kedua model mampu untuk memprediksi zat pencemar CO dikarenakan tingkat error pada hasil prediksi dibawah 20% dan nilai R^2 diatas 0.6.

4.1.3. PM₁₀

1. Data Tanggal 9 Juli 2014 - 23 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 9 Juli 2014 hingga tanggal 22 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verivikasi yaitu pada tanggal 23 Juli 2014. Karakteristik zat pencemar PM₁₀ pada tanggal 9 Juli 2014 hingga tanggal 23 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (23 Juli 2014).

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.19 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|---------|---------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 49.167 | 60.423 | -11.257 | 126.718 | 22.9% |
| 2 | 55.121 | 66.635 | -11.514 | 132.568 | 20.9% |
| 3 | 66.862 | 73.938 | -7.076 | 50.064 | 10.6% |
| 4 | 80.350 | 80.462 | -0.112 | 0.013 | 0.1% |
| 5 | 90.900 | 88.373 | 2.527 | 6.383 | 2.8% |
| 6 | 100.588 | 97.806 | 2.781 | 7.736 | 2.8% |
| 7 | 108.235 | 107.450 | 0.785 | 0.616 | 0.7% |
| 8 | 118.500 | 119.595 | -1.095 | 1.199 | 0.9% |

| | | | | | |
|------------------|-----------------|-----------------|----------------|------------------|--------------|
| 9 | 124.560 | 129.144 | -4.584 | 21.011 | 3.7% |
| 10 | 121.525 | 134.037 | -12.512 | 156.542 | 10.3% |
| 11 | 111.900 | 136.198 | -24.298 | 590.383 | 21.7% |
| 12 | 102.100 | 131.428 | -29.328 | 860.138 | 28.7% |
| 13 | 94.596 | 125.372 | -30.776 | 947.176 | 32.5% |
| 14 | 86.225 | 113.390 | -27.165 | 737.934 | 31.5% |
| 15 | 77.734 | 99.983 | -22.249 | 495.026 | 28.6% |
| 16 | 65.492 | 86.638 | -21.147 | 447.183 | 32.3% |
| 17 | 59.617 | 74.281 | -14.664 | 215.037 | 24.6% |
| 18 | 65.375 | 66.862 | -1.487 | 2.210 | 2.3% |
| 19 | 73.075 | 64.783 | 8.292 | 68.756 | 11.3% |
| 20 | 84.588 | 64.062 | 20.526 | 421.297 | 24.3% |
| 21 | 99.313 | 63.997 | 35.316 | 1247.195 | 35.6% |
| 22 | 116.950 | 65.398 | 51.552 | 2657.586 | 44.1% |
| 23 | 135.175 | 66.746 | 68.429 | 4682.572 | 50.6% |
| Rata-Rata | 90.780 | 92.043 | -1.263 | 603.276 | 19.3% |
| Jumlah | 2087.945 | 2117.000 | -29.055 | 13875.344 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.19 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-11.257}{49.167} \right| \times 100 \%$

$$PE\ 1 = 22.9\ \%$$

$$PE\ 2 = \left| \frac{-11.514}{55.121} \right| \times 100\ \%$$

$$PE\ 2 = 20.9\ \%$$

Dilanjutkan sampai PE ke 23

$$PE\ 23 = \left| \frac{68.429}{135.175} \right| \times 100\ \%$$

$$PE\ 23 = 50.6\ \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{22.9\% + 20.9\% \dots \dots 50.6\%}{23}$$

$$MAPE = 19.3\ \%$$

- R^2

Didapatkan langsung dari software $R = 0.278$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.278} \\ &= 0.527 \end{aligned}$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas

terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.20 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|---------|---------|--------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 49.167 | 58.237 | -9.070 | 82.265 | 18.4% |
| 2 | 55.121 | 63.033 | -7.912 | 62.605 | 14.4% |
| 3 | 66.862 | 68.854 | -1.991 | 3.965 | 3.0% |
| 4 | 80.350 | 73.250 | 7.100 | 50.412 | 8.8% |
| 5 | 90.900 | 77.835 | 13.065 | 170.704 | 14.4% |
| 6 | 100.588 | 85.540 | 15.048 | 226.432 | 15.0% |
| 7 | 108.235 | 93.421 | 14.814 | 219.444 | 13.7% |
| 8 | 118.500 | 102.898 | 15.602 | 243.415 | 13.2% |
| 9 | 124.560 | 110.975 | 13.585 | 184.558 | 10.9% |
| 10 | 121.525 | 112.940 | 8.585 | 73.709 | 7.1% |
| 11 | 111.900 | 111.107 | 0.793 | 0.628 | 0.7% |
| 12 | 102.100 | 106.481 | -4.381 | 19.195 | 4.3% |
| 13 | 94.596 | 100.969 | -6.373 | 40.619 | 6.7% |
| 14 | 86.225 | 88.851 | -2.626 | 6.894 | 3.0% |
| 15 | 77.734 | 76.799 | 0.935 | 0.874 | 1.2% |
| 16 | 65.492 | 65.286 | 0.205 | 0.042 | 0.3% |
| 17 | 59.617 | 54.687 | 4.930 | 24.302 | 8.3% |
| 18 | 65.375 | 50.519 | 14.856 | 220.691 | 22.7% |
| 19 | 73.075 | 51.333 | 21.742 | 472.728 | 29.8% |
| 20 | 84.588 | 51.279 | 33.308 | 1109.428 | 39.4% |

| | | | | | |
|------------------|-----------------|-----------------|----------------|------------------|--------------|
| 21 | 99.313 | 51.659 | 47.654 | 2270.863 | 48.0% |
| 22 | 116.950 | 53.175 | 63.775 | 4067.218 | 54.5% |
| 23 | 135.175 | 54.375 | 80.800 | 6528.712 | 59.8% |
| Rata-Rata | 90.780 | 76.674 | 14.106 | 699.117 | 17.3% |
| Jumlah | 2087.945 | 1763.502 | 324.443 | 16079.700 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.20 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-9.070}{49.167} \right| \times 100 \%$

$$PE\ 1 = 18.4 \%$$

- $PE\ 2 = \left| \frac{-7.912}{55.121} \right| \times 100 \%$

$$PE\ 2 = 14.4 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{80.800}{135.175} \right| \times 100 \%$

$$PE\ 23 = 59.8 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{18.4\% + 14.4\% \dots \dots 59.8\%}{23}$$

$$MAPE = 17.3 \%$$

- **R²**

Didapatkan langsung dari software R = 0.279

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.279} \\ &= 0.528 \end{aligned}$$

Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.21

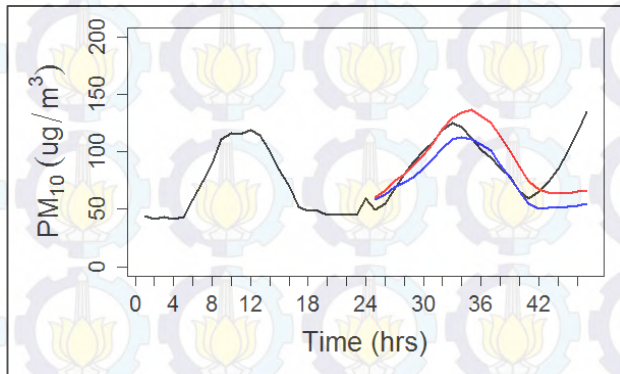
Tabel 4.21 Performa Model ARIMA dan STL

| Data In Sample | | 9 Juli 2014- 23 Juli 2014 | |
|--------------------------------|----------|---------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 19.3% | 17.3% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.278 | 0.279 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.527 | 0.528 |

Berdasarkan Tabel 4.21 performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL dan MAPE kedua model sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R² dan korelasi. Model ARIMA memiliki nilai R² dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.21

nilai R^2 ARIMA sebesar 0.278 sedangkan ARIMA hanya 0.279. Besarnya nilai R^2 dan korelasi pada kedua model tidak memenuhi standard.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.4



Gambar 4. 7 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter PM₁₀

Berdasarkan Gambar 4.4 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada grafik tersebut terlihat kedua model memiliki grafik yang hampir serupa, terbukti karena nilai R^2 kedua model yang tercantum pada Tabel 4.21 tidak terpaut jauh.

Model STL dan ARIMA memiliki nilai R^2 yang tidak sesuai dengan standard yang telah ditetapkan karena dibawah 0.6. Terbukti pada gambar grafik tersebut, pola model STL dan ARIMA tidak mampu mengikuti pola data verifikasi,

Jadi dapat disimpulkan pada pengujian pertama, kedua model tidak mampu untuk memprediksi zat pencemar PM₁₀

dikarenakan nilai R^2 tidak sesuai dengan standard, meskipun tingkat error pada hasil prediksi dibawah 20%.

2. Data Tanggal 12 Juli 2014 - 26 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 12 Juli 2014 hingga tanggal 25 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verivikasi yaitu pada tanggal 26 Juli 2014. Karakteristik zat pencemar PM_{10} pada tanggal 12 Juli 2014 hingga tanggal 25 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (26 Juli 2014). Model terbaik adalah model dengan performa diatas standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model :

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.22 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 30.863 | 26.229 | 4.633 | 21.4688 | 15.0% |
| 2 | 31.363 | 26.672 | 4.691 | 22.0031 | 15.0% |
| 3 | 27.708 | 29.764 | -2.055 | 4.2243 | 7.4% |
| 4 | 25.325 | 32.037 | -6.712 | 45.0481 | 26.5% |
| 5 | 23.616 | 34.573 | -10.956 | 120.0365 | 46.4% |
| 6 | 22.292 | 40.779 | -18.487 | 341.7859 | 82.9% |
| 7 | 23.441 | 46.287 | -22.845 | 521.9083 | 97.5% |
| 8 | 29.934 | 56.599 | -26.666 | 711.0573 | 89.1% |
| 9 | 35.792 | 64.211 | -28.420 | 807.6826 | 79.4% |
| 10 | 43.938 | 67.474 | -23.536 | 553.9643 | 53.6% |

| | | | | | |
|------------------|----------------|----------------|---------------|-----------------|------------|
| 11 | 46.000 | 66.887 | -20.887 | 436.2738 | 45.4% |
| 12 | 46.804 | 61.300 | -14.496 | 210.1469 | 31.0% |
| 13 | 49.450 | 55.244 | -5.794 | 33.5693 | 11.7% |
| 14 | 47.487 | 42.455 | 5.033 | 25.3299 | 10.6% |
| 15 | 46.146 | 28.896 | 17.250 | 297.5732 | 37.4% |
| 16 | 39.066 | 14.845 | 24.221 | 586.6571 | 62.0% |
| 17 | 30.050 | 3.321 | 26.729 | 714.4167 | 88.9% |
| 18 | 23.575 | -3.409 | 26.984 | 728.1396 | 114.5% |
| 19 | 25.750 | -4.812 | 30.562 | 934.0100 | 118.7% |
| 20 | 22.250 | -4.848 | 27.098 | 734.3105 | 121.8% |
| 21 | 21.046 | -3.732 | 24.778 | 613.9327 | 117.7% |
| 22 | 21.967 | -0.670 | 22.637 | 512.4197 | 103.1% |
| 23 | 20.054 | 2.571 | 17.484 | 305.6857 | 87.2% |
| Rata-Rata | 31.909 | 29.681 | 2.228 | 403.550 | 64% |
| Jumlah | 733.917 | 682.673 | 51.244 | 9281.645 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.22 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE_1 = \left| \frac{4.633}{30.863} \right| \times 100 \%$$

$$PE_1 = 15.0 \%$$

- $PE\ 2 = \left| \frac{4.691}{31.363} \right| \times 100\ %$

$$PE\ 2 = 15.0\ %$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{17.484}{20.054} \right| \times 100\ %$

$$PE\ 23 = 7.4\ %$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{15.0\% + 15.0\% \dots \dots 7.4\%}{23}$$

$$MAPE = 64\ %$$

- R^2

Didapatkan langsung dari software $R = 0.412$

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.412}$$

$$= 0.642$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan

PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.23 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|---------|-----------------------|--------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 30.863 | 29.874 | 0.988 | 0.977 | 3.2% |
| 2 | 31.363 | 32.024 | -0.662 | 0.438 | 2.1% |
| 3 | 27.708 | 37.973 | -10.265 | 105.363 | 37.0% |
| 4 | 25.325 | 43.625 | -18.300 | 334.893 | 72.3% |
| 5 | 23.616 | 49.341 | -25.725 | 661.754 | 108.9% |
| 6 | 22.292 | 57.704 | -35.412 | 1254.039 | 158.9% |
| 7 | 23.441 | 65.723 | -42.281 | 1787.708 | 180.4% |
| 8 | 29.934 | 77.728 | -47.795 | 2284.338 | 159.7% |
| 9 | 35.792 | 87.527 | -51.736 | 2676.580 | 144.5% |
| 10 | 43.938 | 92.002 | -48.065 | 2310.226 | 109.4% |
| 11 | 46.000 | 92.470 | -46.470 | 2159.483 | 101.0% |
| 12 | 46.804 | 89.293 | -42.489 | 1805.287 | 90.8% |
| 13 | 49.450 | 84.640 | -35.190 | 1238.324 | 71.2% |
| 14 | 47.487 | 75.639 | -28.151 | 792.484 | 59.3% |
| 15 | 46.146 | 64.666 | -18.520 | 343.004 | 40.1% |
| 16 | 39.066 | 53.458 | -14.391 | 207.109 | 36.8% |
| 17 | 30.050 | 44.936 | -14.886 | 221.588 | 49.5% |
| 18 | 23.575 | 40.666 | -17.091 | 292.112 | 72.5% |
| 19 | 25.750 | 39.788 | -14.038 | 197.067 | 54.5% |
| 20 | 22.250 | 40.349 | -18.099 | 327.561 | 81.3% |
| 21 | 21.046 | 42.425 | -21.379 | 457.070 | 101.6% |
| 22 | 21.967 | 45.266 | -23.299 | 542.866 | 106.1% |
| 23 | 20.054 | 48.970 | -28.916 | 836.134 | 144.2% |

| | | | | | |
|-----------|---------|----------|----------|-----------|-----|
| Rata-Rata | 31.909 | 58.091 | -26.181 | 905.931 | 86% |
| Jumlah | 733.917 | 1336.088 | -602.172 | 20836.406 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.23 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{0.988}{30.863} \right| \times 100 \%$

$$PE\ 1 = 3.2 \%$$

- $PE\ 2 = \left| \frac{-0.662}{31.363} \right| \times 100 \%$

$$PE\ 2 = 2.1 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-28.916}{20.054} \right| \times 100 \%$

$$PE\ 23 = 144.2 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{3.2\% + 2.1\% \dots \dots 144.2\%}{23}$$

$$MAPE = 86 \%$$

- **R²**

Didapatkan langsung dari software $R = 0.520$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.520} \\ &= 0.721 \end{aligned}$$

Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.21

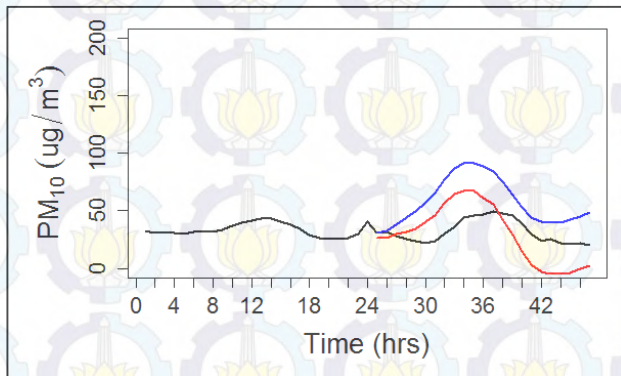
Tabel 4.24 Performa Model ARIMA dan STL

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 64% | 86% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.412 | 0.520 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.642 | 0.721 |

Berdasarkan Tabel 4.24 nilai MAPE pada Model STL dan Model ARIMA lebih besar dari standard yang telah ditetapkan, meskipun performa STL lebih bagus dibandingkan ARIMA karena MAPE pada Model STL lebih kecil daripada ARIMA. Besarnya nilai MAPE disebabkan karena prediksi (P) kedua model banyak yang menyimpang dari data Asli (O). terbukti pada nomor 5 hingga 12, hasil prediksi (P) menyimpang dua kali lipat dari data asli (O), sehingga mengakibatkan nilai |PE| diatas 100%

Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model ARIMA memiliki nilai R^2 dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.24 nilai R^2 ARIMA sebesar 0.520 sedangkan STL hanya 0.412, meskipun begitu besarnya nilai R^2 dan korelasi pada kedua model tidak memenuhi standard.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.8



Gambar 4.8 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter PM₁₀

Berdasarkan Gambar 4.4 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Model STL dan ARIMA memiliki nilai R^2 yang tidak sesuai dengan standard yang telah ditetapkan karena dibawah 0.6. Terbukti pada gambar grafik tersebut, pola model STL dan ARIMA tidak mampu mengikuti pola data verifikasi,

Jadi dapat disimpulkan pada pengujian kedua, kedua model tidak mampu untuk memprediksi zat pencemar PM₁₀ dikarenakan nilai R^2 tidak sesuai dengan standard dan tingkat error pada hasil prediksi lebih besar dari 20%.

3. Data Tanggal 15 Juli 2014 - 29 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 15 Juli 2014 hingga tanggal 28 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 29 Juli 2014. Karakteristik zat pencemar PM_{10} pada tanggal 15 Juli 2014 hingga tanggal 28 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (29 Juli 2014).

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model :

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.25 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|--------|-----------------------|--------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 4.45 | -4.170 | 13.357 | 178.415 | 300.2% |
| 2 | 9.188 | -4.929 | 20.850 | 434.710 | 226.9% |
| 3 | 15.921 | -3.789 | 25.989 | 675.428 | 163.2% |
| 4 | 22.200 | -1.823 | 30.123 | 907.380 | 135.7% |
| 5 | 28.300 | 0.219 | 35.948 | 1292.226 | 127.0% |
| 6 | 36.167 | 4.823 | 37.652 | 1417.667 | 104.1% |
| 7 | 42.475 | 9.720 | 39.842 | 1587.396 | 93.8% |
| 8 | 49.563 | 17.644 | 37.085 | 1375.270 | 74.8% |
| 9 | 54.729 | 25.236 | 31.697 | 1004.711 | 57.9% |
| 10 | 56.934 | 28.084 | 23.999 | 575.957 | 42.2% |
| 11 | 52.083 | 25.838 | 20.112 | 404.498 | 38.6% |
| 12 | 45.950 | 19.417 | 20.316 | 412.745 | 44.2% |
| 13 | 39.733 | 13.650 | 18.121 | 328.361 | 45.6% |
| 14 | 31.771 | 2.275 | 24.287 | 589.859 | 76.4% |

| | | | | | |
|------------------|----------------|-----------------|----------------|------------------|-------------|
| 15 | 26.563 | -10.376 | 30.297 | 917.933 | 114.1% |
| 16 | 19.921 | -23.999 | 42.357 | 1794.150 | 212.6% |
| 17 | 18.358 | -35.265 | 47.040 | 2212.786 | 256.2% |
| 18 | 11.775 | -41.495 | 51.753 | 2678.379 | 439.5% |
| 19 | 10.259 | -43.833 | 55.979 | 3133.629 | 545.7% |
| 20 | 12.146 | -45.500 | 58.825 | 3460.383 | 484.3% |
| 21 | 13.325 | -45.262 | 59.724 | 3566.957 | 448.2% |
| 22 | 14.463 | -42.489 | 55.255 | 3053.158 | 382.1% |
| 23 | 12.766 | -39.768 | 67.117 | 4504.692 | 525.7% |
| Rata-Rata | 27.349 | -8.513 | 36.858 | 1587.247 | 215% |
| Jumlah | 629.038 | -195.788 | 847.726 | 36506.691 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.25 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{13.357}{4.45} \right| \times 100 \%$

$$PE\ 1 = 300.2 \%$$

- $PE\ 2 = \left| \frac{20.850}{9.188} \right| \times 100 \%$

$$PE\ 2 = 226.9 \%$$

Dilanjutkan sampai PE ke 23

- $PE_{23} = \left| \frac{67.117}{12.766} \right| \times 100 \%$

$$PE_{23} = 525.7 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{300.2\% + 226.9\% \dots \dots 525.7\%}{23}$$

$$MAPE = 215 \%$$

- R^2

Didapatkan langsung dari software $R = 0.704$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.704} \\ &= 0.839 \end{aligned}$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan

kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.26 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|---------|---------|---------|-----------------------|--------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 4.45 | 3.544 | 0.906 | 0.822 | 20.4% |
| 2 | 9.188 | 5.013 | 4.174 | 17.423 | 45.4% |
| 3 | 15.921 | 6.721 | 9.200 | 84.646 | 57.8% |
| 4 | 22.200 | 9.940 | 12.260 | 150.304 | 55.2% |
| 5 | 28.300 | 12.859 | 15.441 | 238.437 | 54.6% |
| 6 | 36.167 | 15.450 | 20.717 | 429.187 | 57.3% |
| 7 | 42.475 | 18.985 | 23.490 | 551.764 | 55.3% |
| 8 | 49.563 | 22.548 | 27.014 | 729.776 | 54.5% |
| 9 | 54.729 | 26.106 | 28.623 | 819.266 | 52.3% |
| 10 | 56.934 | 23.460 | 33.474 | 1120.484 | 58.8% |
| 11 | 52.083 | 20.762 | 31.321 | 981.009 | 60.1% |
| 12 | 45.950 | 16.413 | 29.537 | 872.419 | 64.3% |
| 13 | 39.733 | 11.497 | 28.236 | 797.298 | 71.1% |
| 14 | 31.771 | 6.201 | 25.570 | 653.842 | 80.5% |
| 15 | 26.563 | 0.579 | 25.984 | 675.160 | 97.8% |
| 16 | 19.921 | -6.079 | 26.000 | 675.985 | 130.5% |
| 17 | 18.358 | -11.106 | 29.464 | 868.146 | 160.5% |
| 18 | 11.775 | -13.321 | 25.096 | 629.806 | 213.1% |
| 19 | 10.259 | -13.717 | 23.975 | 574.801 | 233.7% |
| 20 | 12.146 | -13.985 | 26.131 | 682.850 | 215.1% |
| 21 | 13.325 | -13.363 | 26.688 | 712.251 | 200.3% |
| 22 | 14.463 | -12.383 | 26.845 | 720.657 | 185.6% |
| 23 | 12.766 | -11.258 | 24.024 | 577.157 | 188.2% |
| Rata-Rata | 27.349 | 4.559 | 22.790 | 589.717 | 105% |
| Jumlah | 629.038 | 104.867 | 524.171 | 13563.490 | - |

Keterangan :

- O = Data Asli (data sebenarnya di hari ke-15)
 P = Data hasil prediksi model
 $O - P$ = Kesalahan Prediksi
 n = Jumlah data Pengamatan

Dari Tabel 4.26 dapat diketahui

• **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{0.906}{4.45} \right| \times 100 \%$$

$$PE 1 = 20.4 \%$$

$$\blacksquare PE 2 = \left| \frac{4.174}{9.188} \right| \times 100 \%$$

$$PE 2 = 45.4 \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare PE 23 = \left| \frac{24.024}{12.766} \right| \times 100 \%$$

$$PE 23 = 188.2 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{20.4\% + 45.4\% \dots \dots 188.2\%}{23}$$

$$MAPE = 105 \%$$

- **R²**

Didapatkan langsung dari software $R = 0.733$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.733} \\ &= 0.856 \end{aligned}$$

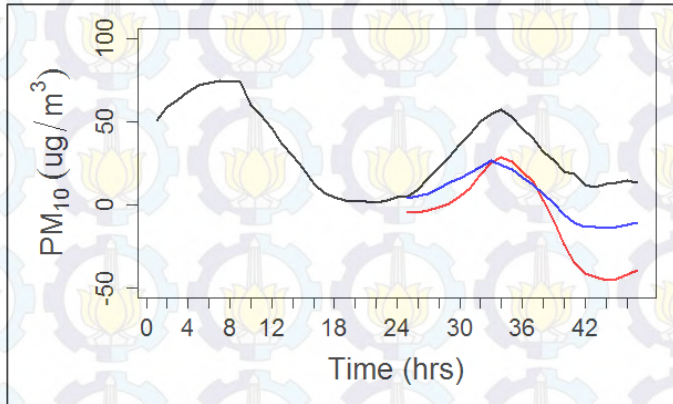
Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.21

Tabel 4.27 Performa Model ARIMA dan STL

| Data In Sample | | 15 Juli 2014 - 29 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 215% | 105% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.124 | 0.733 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.352 | 0.856 |

Berdasarkan Tabel 4.27 nilai MAPE pada Model STL dan Model ARIMA sangat besar dari standard yang telah ditetapkan, meskipun performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL. Besarnya nilai MAPE disebabkan karena prediksi (P) kedua model banyak yang menyimpang dari data Asli (O). terbukti pada Tabel 4.25 dan 4.26 hasil prediksi (P) banyak menyimpang hingga dua kali lipat dari data asli (O), sehingga mengakibatkan nilai |PE| banyak yang diatas 100%. Penilaian performa model juga dibuktikan dengan nilai R² dan korelasi. Model ARIMA memiliki nilai R² dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.27 nilai R² ARIMA sebesar 0.733 sedangkan STL sebesar 0.704.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.9



Gambar 4.9 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter PM₁₀

Berdasarkan Gambar 4.4 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Model STL dan ARIMA memiliki nilai R^2 yang sesuai dengan standard yang telah ditetapkan. Terbukti pada gambar grafik tersebut, pola model STL dan ARIMA mampu mengikuti pola data verifikasi, hanya saja tingkat error pada kedua model yang sangat besar dengan rata-rata diatas 100% mengakibatkan jarak grafik prediksi dengan data verifikasi pada Gambar 4.9 terlihat jauh.

Jadi dapat disimpulkan pada ketiga pengujian kedua model tidak konsisten dalam memprediksi zat pencemar PM₁₀ dan tingkat error banyak yang tidak memenuhi standard, bahkan ada yang melebihi 100%.

4.1.4. O₃

1. Data Tanggal 13 Juli 2014 - 27 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 13 Juli 2014 hingga tanggal 26 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 27 Juli 2014. Karakteristik zat pencemar O₃ pada tanggal 13 Juli 2014 hingga tanggal 26 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (27 Juli 2014)

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R² dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.28 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 26.040 | 31.939 | -5.899 | 34.802 | 22.7% |
| 2 | 26.752 | 33.867 | -7.116 | 50.637 | 26.6% |
| 3 | 21.913 | 34.620 | -12.707 | 161.466 | 58.0% |
| 4 | 21.717 | 36.631 | -14.914 | 222.435 | 68.7% |
| 5 | 22.368 | 37.562 | -15.195 | 230.884 | 67.9% |
| 6 | 19.707 | 38.936 | -19.229 | 369.770 | 97.6% |
| 7 | 21.480 | 41.840 | -20.360 | 414.513 | 94.8% |
| 8 | 31.063 | 50.612 | -19.549 | 382.158 | 62.9% |
| 9 | 51.104 | 65.958 | -14.853 | 220.613 | 29.1% |
| 10 | 58.385 | 86.216 | -27.830 | 774.516 | 47.7% |

| | | | | | |
|------------------|----------------|------------------|------------------|-------------------|------------|
| 11 | 53.747 | 97.048 | -43.302 | 1875.061 | 80.6% |
| 12 | 62.200 | 101.265 | -39.066 | 1526.126 | 62.8% |
| 13 | 66.631 | 89.053 | -22.422 | 502.726 | 33.7% |
| 14 | 59.014 | 82.984 | -23.969 | 574.520 | 40.6% |
| 15 | 55.967 | 80.618 | -24.651 | 607.693 | 44.0% |
| 16 | 52.674 | 74.865 | -22.191 | 492.440 | 42.1% |
| 17 | 45.674 | 72.657 | -26.983 | 728.097 | 59.1% |
| 18 | 50.785 | 71.309 | -20.524 | 421.231 | 40.4% |
| 19 | 47.863 | 74.779 | -26.916 | 724.446 | 56.2% |
| 20 | 38.676 | 78.320 | -39.644 | 1571.626 | 102.5% |
| 21 | 41.192 | 82.861 | -41.669 | 1736.308 | 101.2% |
| 22 | 32.503 | 86.476 | -53.973 | 2913.108 | 166.1% |
| 23 | 26.382 | 88.388 | -62.007 | 3844.817 | 235.0% |
| Rata-Rata | 40.602 | 66.904 | - 26.303 | 886.087 | 71% |
| Jumlah | 933.835 | 1,538.803 | - 604.968 | 20,379.993 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.28 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-5.899}{26.040} \right| \times 100\ %$

$$PE\ 1 = 22.7\ %$$

- $PE\ 2 = \left| \frac{-7.116}{26.752} \right| \times 100\ %$

$$PE\ 2 = 22.6\ %$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-62.007}{26.382} \right| \times 100\ %$

$$PE\ 23 = 235.0\ %$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{22.7\% + 22.6\% \dots \dots 235.0\%}{23}$$

$$MAPE = 71.3\ %$$

- **R²**

Didapatkan langsung dari software R = 0.63

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.63}$$

$$= 0.794$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R² dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap

mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.29 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 26.040 | 29.933 | -3.894 | 15.161 | 15.0% |
| 2 | 26.752 | 37.221 | -10.469 | 109.608 | 39.1% |
| 3 | 21.913 | 38.782 | -16.869 | 284.550 | 77.0% |
| 4 | 21.717 | 39.853 | -18.137 | 328.950 | 83.5% |
| 5 | 22.368 | 38.603 | -16.235 | 263.576 | 72.6% |
| 6 | 19.707 | 38.074 | -18.367 | 337.351 | 93.2% |
| 7 | 21.480 | 38.434 | -16.954 | 287.426 | 78.9% |
| 8 | 31.063 | 51.113 | -20.049 | 401.974 | 64.5% |
| 9 | 51.104 | 46.744 | 4.360 | 19.010 | 8.5% |
| 10 | 58.385 | 48.843 | 9.542 | 91.052 | 16.3% |
| 11 | 53.747 | 48.107 | 5.640 | 31.805 | 10.5% |
| 12 | 62.200 | 47.341 | 14.858 | 220.767 | 23.9% |
| 13 | 66.631 | 47.556 | 19.075 | 363.852 | 28.6% |
| 14 | 59.014 | 46.665 | 12.349 | 152.505 | 20.9% |
| 15 | 55.967 | 45.878 | 10.089 | 101.788 | 18.0% |

| | | | | | |
|------------------|----------------|----------------|----------------|------------------|--------------|
| 16 | 52.674 | 42.591 | 10.083 | 101.663 | 19.1% |
| 17 | 45.674 | 39.050 | 6.624 | 43.873 | 14.5% |
| 18 | 50.785 | 36.424 | 14.362 | 206.266 | 28.3% |
| 19 | 47.863 | 35.397 | 12.466 | 155.411 | 26.0% |
| 20 | 38.676 | 34.619 | 4.057 | 16.463 | 10.5% |
| 21 | 41.192 | 36.921 | 4.271 | 18.240 | 10.4% |
| 22 | 32.503 | 36.866 | -4.363 | 19.037 | 13.4% |
| 23 | 26.382 | 36.308 | -9.927 | 98.544 | 37.6% |
| Rata-Rata | 40.602 | 40.927 | - 0.326 | 159.516 | 35.2% |
| Jumlah | 933.835 | 941.322 | - 7.487 | 3,668.871 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.29 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-3.894}{26.040} \right| \times 100 \%$

$$PE\ 1 = 15.0 \%$$

- $PE\ 2 = \left| \frac{-10.469}{26.752} \right| \times 100 \%$

$$PE\ 2 = 39.1 \%$$

Dilanjutkan sampai PE ke 23

$$PE_{23} = \left| \frac{-9.927}{26.382} \right| \times 100 \%$$

$$PE_{23} = 37.6.0 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{15.0\% + 39.1\% \dots \dots 37.6\%}{23}$$

$$MAPE = 35.2 \%$$

- R^2

Didapatkan langsung dari software $R = 0.35$

- Korelasi

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.35} \\ &= 0.596 \end{aligned}$$

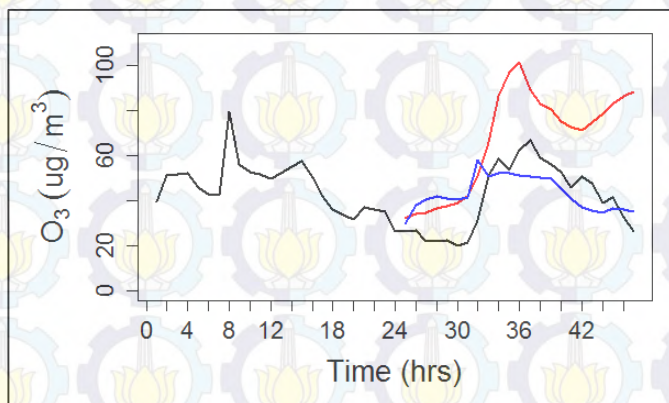
Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.30

Tabel 4.30 Performa Model ARIMA dan STL

| Data In Sample | | 13 Juli 2014 - 27 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 71% | 35% |
| R^2 | STANDARD | > 0.6 | |
| | HASIL | 0.63 | 0.355 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.794 | 0.596 |

Berdasarkan Tabel 4.30 performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL dan MAPE kedua model tidak sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model STL memiliki nilai R^2 dan korelasi yang lebih besar daripada ARIMA, terlihat pada Tabel 4.30 nilai R^2 STL sebesar 0.63 sedangkan ARIMA hanya 0.355. Besarnya nilai R^2 dan korelasi pada kedua model tidak memenuhi standard.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.10



Gambar 4. 10 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter O_3

Berdasarkan Gambar 4.10 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada Grafik tersebut dapat dilihat pola Model STL lebih identik dengan data Asli daripada model ARIMA, tetapi Model STL berjarak cukup jauh dengan data asli.

Hal itu dibuktikan dengan nilai R^2 pada Model STL sebesar 0,63 dan korelasi sebesar 0.794, sedangkan model ARIMA hanya memiliki nilai R^2 0.355 dan nilai korelasi sebesar 0,536. Meskipun nilai R^2 dan korelasi Model STL sudah sesuai standard, akan tetapi tingkat kesalahan prediksi sangat besar, yakni 71.3%.

Jadi dapat disimpulkan pada pengujian pertama, kedua model tidak mampu untuk memprediksi zat pencemar O_3 dikarenakan tidak adanya nilai MAPE dan R^2 pada kedua model tidak konsisten

2. Data Tanggal 9 Juli 2014 - 23 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 9 Juli 2014 hingga tanggal 22 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 23 Juli 2014. Karakteristik zat pencemar PM_{10} pada tanggal 9 Juli 2014 hingga tanggal 23 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (23 Juli 2014).

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.31 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|--------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 44.001 | 43.244 | 0.758 | 0.574 | 1.7% |
| 2 | 37.223 | 44.083 | -6.860 | 47.058 | 18.4% |
| 3 | 35.904 | 42.616 | -6.712 | 45.045 | 18.7% |

| | | | | | |
|------------------|----------------|------------------|------------------|------------------|------------|
| 4 | 33.511 | 42.725 | -9.214 | 84.893 | 27.5% |
| 5 | 28.647 | 42.039 | -13.391 | 179.331 | 46.7% |
| 6 | 23.854 | 41.755 | -17.901 | 320.445 | 75.0% |
| 7 | 25.153 | 43.230 | -18.077 | 326.779 | 71.9% |
| 8 | 33.390 | 50.975 | -17.585 | 309.226 | 52.7% |
| 9 | 50.160 | 65.633 | -15.473 | 239.421 | 30.8% |
| 10 | 51.809 | 81.871 | -30.062 | 903.742 | 58.0% |
| 11 | 52.550 | 86.546 | -33.996 | 1155.705 | 64.7% |
| 12 | 57.014 | 86.405 | -29.391 | 863.844 | 51.6% |
| 13 | 62.583 | 80.536 | -17.953 | 322.323 | 28.7% |
| 14 | 57.426 | 73.487 | -16.061 | 257.961 | 28.0% |
| 15 | 46.203 | 69.658 | -23.456 | 550.168 | 50.8% |
| 16 | 44.285 | 65.074 | -20.789 | 432.185 | 46.9% |
| 17 | 43.577 | 61.320 | -17.743 | 314.803 | 40.7% |
| 18 | 48.107 | 58.606 | -10.500 | 110.242 | 21.8% |
| 19 | 45.538 | 60.269 | -14.730 | 216.979 | 32.3% |
| 20 | 43.879 | 61.261 | -17.382 | 302.146 | 39.6% |
| 21 | 38.745 | 64.235 | -25.490 | 649.736 | 65.8% |
| 22 | 36.054 | 65.865 | -29.811 | 888.697 | 82.7% |
| 23 | 39.330 | 65.832 | -26.502 | 702.334 | 67.4% |
| Rata-Rata | 42.563 | 60.751 | - 18.188 | 401.028 | 44% |
| Jumlah | 978.943 | 1,397.263 | - 418.321 | 9,223.639 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.31 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{0.758}{44.001} \right| \times 100 \%$

$$PE\ 1 = 1.7 \%$$

- $PE\ 2 = \left| \frac{-6.860}{37.223} \right| \times 100 \%$

$$PE\ 2 = 18.4 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-26.502}{39.330} \right| \times 100 \%$

$$PE\ 23 = 67.4 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{1.7\% + 18.4\% \dots \dots 67.4\%}{23}$$

$$MAPE = 44 \%$$

- **R²**

Didapatkan langsung dari software R = 0.689

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.689} \end{aligned}$$

$$= 0.830$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah `auto.arima`, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.32 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|---------|---------|--------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 49.167 | 58.237 | -9.070 | 82.265 | 18.4% |
| 2 | 55.121 | 63.033 | -7.912 | 62.605 | 14.4% |
| 3 | 66.862 | 68.854 | -1.991 | 3.965 | 3.0% |
| 4 | 80.350 | 73.250 | 7.100 | 50.412 | 8.8% |
| 5 | 90.900 | 77.835 | 13.065 | 170.704 | 14.4% |
| 6 | 100.588 | 85.540 | 15.048 | 226.432 | 15.0% |
| 7 | 108.235 | 93.421 | 14.814 | 219.444 | 13.7% |
| 8 | 118.500 | 102.898 | 15.602 | 243.415 | 13.2% |
| 9 | 124.560 | 110.975 | 13.585 | 184.558 | 10.9% |
| 10 | 121.525 | 112.940 | 8.585 | 73.709 | 7.1% |

| | | | | | |
|-----------|----------|----------|---------|-----------|-------|
| 11 | 111.900 | 111.107 | 0.793 | 0.628 | 0.7% |
| 12 | 102.100 | 106.481 | -4.381 | 19.195 | 4.3% |
| 13 | 94.596 | 100.969 | -6.373 | 40.619 | 6.7% |
| 14 | 86.225 | 88.851 | -2.626 | 6.894 | 3.0% |
| 15 | 77.734 | 76.799 | 0.935 | 0.874 | 1.2% |
| 16 | 65.492 | 65.286 | 0.205 | 0.042 | 0.3% |
| 17 | 59.617 | 54.687 | 4.930 | 24.302 | 8.3% |
| 18 | 65.375 | 50.519 | 14.856 | 220.691 | 22.7% |
| 19 | 73.075 | 51.333 | 21.742 | 472.728 | 29.8% |
| 20 | 84.588 | 51.279 | 33.308 | 1109.428 | 39.4% |
| 21 | 99.313 | 51.659 | 47.654 | 2270.863 | 48.0% |
| 22 | 116.950 | 53.175 | 63.775 | 4067.218 | 54.5% |
| 23 | 135.175 | 54.375 | 80.800 | 6528.712 | 59.8% |
| Rata-Rata | 90.780 | 76.674 | 14.106 | 699.117 | 17.3% |
| Jumlah | 2087.945 | 1763.502 | 324.443 | 16079.700 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.32 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{3.976}{44.001} \right| \times 100 \%$$

$$PE 1 = 9.0 \%$$

- $PE\ 2 = \left| \frac{-1.423}{37.223} \right| \times 100\ %$

$$PE\ 2 = 3.8\ %$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-0.114}{39.330} \right| \times 100\ %$

$$PE\ 23 = 0.3\ %$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{9\% + 3.8\% \dots \dots 0.3\%}{23}$$

$$MAPE = 13\ %$$

- R^2

Didapatkan langsung dari software $R = 0.837$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.837} \\ &= 0.915 \end{aligned}$$

Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.33

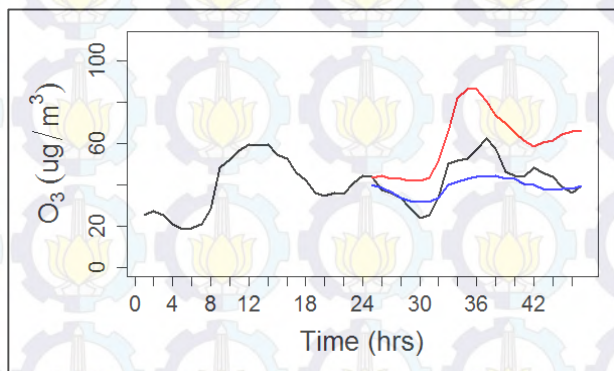
Tabel 4.33 Performa Model ARIMA dan STL

| Data In Sample | | 16 Juli 2014 - 30 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 44% | 13% |
| R^2 | STANDARD | > 0.6 | |

| | | | |
|-------------|----------|------------|-------|
| Correlation | HASIL | 0.689 | 0.837 |
| | STANDARD | (+ 0.774) | |
| | HASIL | 0.830 | 0.915 |

Berdasarkan Tabel 4.33 performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL dan hanya MAPE model ARIMA yang sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model ARIMA memiliki nilai R^2 dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.33 nilai R^2 ARIMA sebesar 0.837 sedangkan ARIMA hanya 0.689. Besarnya nilai R^2 dan korelasi pada kedua telah memenuhi standard.

Nilai R^2 dan korelasi menunjukan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.11



Gambar 4.11 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter O_3

Berdasarkan Gambar 4.11 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam

adalah data verifikasi (data asli). Model STL dan ARIMA memiliki nilai R^2 yang sesuai dengan standard yang telah ditetapkan karena diatas 0.6. Terbukti pada gambar grafik tersebut, pola model STL dan ARIMA cukup mampu mengikuti pola data verifikasi, tingkat error pada Model STL yang besar mengakibatkan grafik hasil prediksi model tersebut jauh dari data asli

Jadi dapat disimpulkan pada pengujian kedua, model ARIMA mampu untuk memprediksi zat pencemar O_3 dikarenakan nilai R^2 yang telah sesuai dengan standard, dan tingkat error yang dbawah 20%.

3. Data Tanggal 19 Juli 2014 - 31 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 19 Juli 2014 hingga tanggal 30 Juli 2014 (14 hari) dan satu hari digunakan sebagai data verivikasi yaitu pada tanggal 31 Juli 2014. Karakteristik zat pencemar O_3 pada tanggal 19 Juli 2014 hingga tanggal 30 Juli 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (31 Juli 2014)

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.34 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 30.454 | 39.734 | -9.280 | 86.115 | 30.5% |
| 2 | 25.991 | 40.417 | -14.427 | 208.127 | 55.5% |

| | | | | | |
|------------------|----------------|------------------|------------------|------------------|------------|
| 3 | 28.624 | 38.330 | -9.706 | 94.203 | 33.9% |
| 4 | 27.095 | 36.772 | -9.677 | 93.639 | 35.7% |
| 5 | 28.424 | 35.699 | -7.275 | 52.924 | 25.6% |
| 6 | 26.564 | 35.049 | -8.484 | 71.982 | 31.9% |
| 7 | 26.403 | 36.151 | -9.748 | 95.028 | 36.9% |
| 8 | 31.228 | 44.550 | -13.322 | 177.472 | 42.7% |
| 9 | 45.629 | 58.293 | -12.664 | 160.365 | 27.8% |
| 10 | 73.097 | 67.996 | 5.101 | 26.016 | 7.0% |
| 11 | 68.875 | 66.611 | 2.263 | 5.121 | 3.3% |
| 12 | 65.524 | 69.017 | -3.492 | 12.194 | 5.3% |
| 13 | 68.766 | 69.226 | -0.460 | 0.211 | 0.7% |
| 14 | 60.957 | 63.884 | -2.927 | 8.567 | 4.8% |
| 15 | 49.514 | 58.522 | -9.009 | 81.161 | 18.2% |
| 16 | 37.711 | 56.914 | -19.203 | 368.760 | 50.9% |
| 17 | 37.755 | 54.492 | -16.737 | 280.126 | 44.3% |
| 18 | 45.771 | 53.242 | -7.471 | 55.817 | 16.3% |
| 19 | 45.956 | 55.562 | -9.606 | 92.273 | 20.9% |
| 20 | 37.439 | 55.666 | -18.227 | 332.215 | 48.7% |
| 21 | 38.202 | 57.972 | -19.770 | 390.834 | 51.7% |
| 22 | 42.276 | 58.482 | -16.206 | 262.634 | 38.3% |
| 23 | 33.219 | 59.392 | -26.173 | 685.013 | 78.8% |
| Rata-Rata | 42.412 | 52.694 | - 10.283 | 158.296 | 31% |
| Jumlah | 975.475 | 1,211.973 | - 236.498 | 3,640.798 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

O - P = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.34 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-9.280}{30.454} \right| \times 100 \%$
 $PE\ 1 = 30.5 \%$

- $PE\ 2 = \left| \frac{-14.427}{25.991} \right| \times 100 \%$
 $PE\ 2 = 55.5 \%$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-26.173}{33.219} \right| \times 100 \%$
 $PE\ 23 = 78.8 \%$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{30.5\% + 55.5\% \dots \dots 78.8\%}{23}$$

$$MAPE = 31 \%$$

- **R²**

Didapatkan langsung dari software $R = 0.778$

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.778}$$

$$= 0.882$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.35 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 30.454 | 38.721 | -8.267 | 68.339 | 27.1% |
| 2 | 25.991 | 37.796 | -11.805 | 139.359 | 45.4% |
| 3 | 28.624 | 34.717 | -6.093 | 37.125 | 21.3% |
| 4 | 27.095 | 31.622 | -4.526 | 20.485 | 16.7% |
| 5 | 28.424 | 30.002 | -1.578 | 2.491 | 5.6% |
| 6 | 26.564 | 29.301 | -2.737 | 7.490 | 10.3% |
| 7 | 26.403 | 29.738 | -3.335 | 11.122 | 12.6% |

| | | | | | |
|------------------|----------------|----------------|---------------|------------------|------------|
| 8 | 31.228 | 35.922 | -4.694 | 22.030 | 15.0% |
| 9 | 45.629 | 46.932 | -1.303 | 1.698 | 2.9% |
| 10 | 73.097 | 51.667 | 21.429 | 459.212 | 29.3% |
| 11 | 68.875 | 50.741 | 18.134 | 328.839 | 26.3% |
| 12 | 65.524 | 51.533 | 13.991 | 195.756 | 21.4% |
| 13 | 68.766 | 52.543 | 16.224 | 263.208 | 23.6% |
| 14 | 60.957 | 49.949 | 11.008 | 121.186 | 18.1% |
| 15 | 49.514 | 45.683 | 3.831 | 14.675 | 7.7% |
| 16 | 37.711 | 44.286 | -6.576 | 43.238 | 17.4% |
| 17 | 37.755 | 40.741 | -2.986 | 8.913 | 7.9% |
| 18 | 45.771 | 40.398 | 5.373 | 28.869 | 11.7% |
| 19 | 45.956 | 41.486 | 4.471 | 19.986 | 9.7% |
| 20 | 37.439 | 40.668 | -3.229 | 10.429 | 8.6% |
| 21 | 38.202 | 42.108 | -3.905 | 15.251 | 10.2% |
| 22 | 42.276 | 40.598 | 1.678 | 2.816 | 4.0% |
| 23 | 33.219 | 41.292 | -8.073 | 65.167 | 24.3% |
| Rata-Rata | 42.412 | 1.237 | 1.175 | 82.073 | 16% |
| Jumlah | 975.475 | 948.442 | 27.033 | 1,887.684 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.35 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{-8.267}{30.454} \right| \times 100 \%$$

$$PE 1 = 27.1 \%$$

$$\blacksquare PE 2 = \left| \frac{-11.805}{25.991} \right| \times 100 \%$$

$$PE 2 = 45.4 \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare PE 23 = \left| \frac{-8.073}{33.219} \right| \times 100 \%$$

$$PE 23 = 24.3 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{27.1\% + 45.4\% \dots \dots 24.3\%}{23}$$

$$MAPE = 16 \%$$

- **R²**

Didapatkan langsung dari software R = 0.829

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.829} \\ &= 0.910 \end{aligned}$$

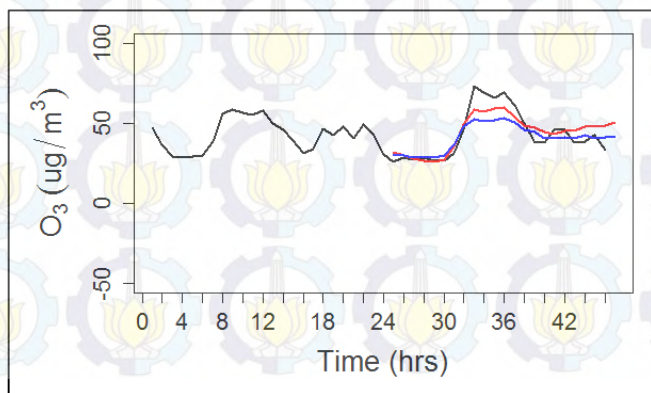
Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.33

Tabel 4.36 Performa Model ARIMA dan STL

| Data In Sample | | 19 Juli 2014 - 31 Juli 2014 | |
|--------------------------------|----------|-----------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 31% | 16% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.778 | 0.829 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.882 | 0.910 |

Berdasarkan Tabel 4.36 performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL dan hanya MAPE model ARIMA yang sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model ARIMA memiliki nilai R^2 dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.33 nilai R^2 ARIMA sebesar 0.829 sedangkan STL hanya 0.778. Besarnya nilai R^2 dan korelasi pada kedua telah memenuhi standard.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.12



Gambar 4.12 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter O_3

Berdasarkan Gambar 4.12 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Model STL dan ARIMA memiliki nilai R^2 yang sesuai dengan standard yang telah ditetapkan karena diatas 0.6. Terbukti pada gambar grafik tersebut, pola model STL dan ARIMA cukup mampu mengikuti pola data verifikasi, tingkat error pada Model STL yang besar mengakibatkan grafik hasil prediksi model tersebut jauh dari data asli

Jadi dapat disimpulkan pada pengujian kedua, model ARIMA mampu untuk memprediksi zat pencemar O_3 dikarenakan nilai R^2 yang telah sesuai dengan standard, dan tingkat error yang dbawah 20%.

4.1.5. NO_2

1. Data Tanggal 29 Juli 2014 - 12 Agustus 2014

Pengujian model menggunakan data in sample pada tanggal 29 Juli 2014 hingga tanggal 11 Agustus 2014 (14 hari) dan satu hari digunakan sebagai data verivikasi yaitu pada tanggal 12 Agustus 2014. Karakteristik zat pencemar NO_2 pada tanggal 29 Juli 2014 hingga tanggal 11 Agustus 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (12 Agustus 2014).

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.37 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|-----------|---------|----------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 16.378 | 31.705 | - 15.328 | 234.946 | 93.6% |
| 2 | 21.811 | 30.691 | - 8.881 | 78.867 | 40.7% |
| 3 | 33.323 | 28.508 | 4.815 | 23.181 | 14.4% |
| 4 | 32.742 | 29.306 | 3.436 | 11.805 | 10.5% |
| 5 | 34.598 | 29.642 | 4.955 | 24.552 | 14.3% |
| 6 | 34.182 | 29.963 | 4.218 | 17.794 | 12.3% |
| 7 | 42.125 | 35.248 | 6.877 | 47.289 | 16.3% |
| 8 | 69.986 | 44.703 | 25.282 | 639.191 | 36.1% |
| 9 | 57.768 | 41.142 | 16.625 | 276.399 | 28.8% |
| 10 | 43.566 | 35.989 | 7.576 | 57.398 | 17.4% |
| 11 | 39.643 | 32.246 | 7.396 | 54.706 | 18.7% |
| 12 | 37.982 | 32.160 | 5.821 | 33.885 | 15.3% |
| 13 | 38.960 | 29.400 | 9.560 | 91.402 | 24.5% |
| 14 | 39.425 | 30.797 | 8.628 | 74.439 | 21.9% |
| 15 | 51.162 | 33.073 | 18.088 | 327.183 | 35.4% |
| 16 | 57.420 | 37.270 | 20.150 | 406.018 | 35.1% |
| 17 | 74.766 | 42.651 | 32.115 | 1,031.349 | 43.0% |
| 18 | 74.962 | 46.150 | 28.812 | 830.132 | 38.4% |
| 19 | 63.192 | 44.974 | 18.218 | 331.885 | 28.8% |
| 20 | 61.308 | 46.256 | 15.052 | 226.556 | 24.6% |
| 21 | 47.187 | 45.878 | 1.309 | 1.715 | 2.8% |
| 22 | 40.060 | 46.279 | - 6.219 | 38.676 | 15.5% |
| 23 | 41.117 | 46.192 | - 5.075 | 25.759 | 12.3% |
| Rata-Rata | 45.811 | 36.966 | 8.845 | 212.397 | 26% |
| Jumlah | 1,053.656 | 850.225 | 203.431 | 4,885.128 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.37 dapat diketahui

• **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{-15.328}{16.378} \right| \times 100 \%$$

$$PE 1 = 93.6 \%$$

$$\blacksquare PE 2 = \left| \frac{-8.881}{21.811} \right| \times 100 \%$$

$$PE 2 = 40.7 \%$$

Dilanjutkan sampai PE ke 23

$$\blacksquare PE 23 = \left| \frac{-5.075}{41.117} \right| \times 100 \%$$

$$PE 23 = 12.3 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{93.6\% + 40.7\% \dots \dots 12.3\%}{23}$$

$$MAPE = 26 \%$$

- R^2

Didapatkan langsung dari software $R = 0.501$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.501} \\ &= 0.708 \end{aligned}$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.38 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|----------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 16.378 | 29.79653 | -13.419 | 180.0703661 | 81.9% |
| 2 | 21.811 | 34.02419 | -12.214 | 149.1742234 | 56.0% |
| 3 | 33.323 | 32.13858 | 1.184 | 1.401666566 | 3.6% |

| | | | | | |
|------------------|------------------|----------------|----------------|------------------|------------|
| 4 | 32.742 | 32.82584 | -0.084 | 0.007113236 | 0.3% |
| 5 | 34.598 | 31.73101 | 2.866 | 8.21676492 | 8.3% |
| 6 | 34.182 | 33.25077 | 0.931 | 0.866258333 | 2.7% |
| 7 | 42.125 | 37.55285 | 4.572 | 20.90455562 | 10.9% |
| 8 | 69.986 | 46.68306 | 23.302 | 543.00371 | 33.3% |
| 9 | 57.768 | 41.59529 | 16.172 | 261.5403763 | 28.0% |
| 10 | 43.566 | 39.43191 | 4.134 | 17.08656629 | 9.5% |
| 11 | 39.643 | 36.75793 | 2.885 | 8.320744085 | 7.3% |
| 12 | 37.982 | 36.66492 | 1.317 | 1.733382896 | 3.5% |
| 13 | 38.960 | 34.03871 | 4.921 | 24.21909526 | 12.6% |
| 14 | 39.425 | 35.08349 | 4.341 | 18.84436782 | 11.0% |
| 15 | 51.162 | 37.74532 | 13.416 | 179.9938858 | 26.2% |
| 16 | 57.420 | 41.48378 | 15.936 | 253.9631079 | 27.8% |
| 17 | 74.766 | 46.84217 | 27.923 | 779.7123583 | 37.3% |
| 18 | 74.962 | 45.79758 | 29.164 | 850.5633939 | 38.9% |
| 19 | 63.192 | 42.72317 | 20.469 | 418.9730016 | 32.4% |
| 20 | 61.308 | 44.55293 | 16.755 | 280.7323707 | 27.3% |
| 21 | 47.187 | 44.54274 | 2.644 | 6.992110948 | 5.6% |
| 22 | 40.060 | 43.60687 | -3.547 | 12.58383392 | 8.9% |
| 23 | 41.117 | 43.0065 | -1.890 | 3.57021025 | 4.6% |
| Rata-Rata | 45.811 | 38.777 | 7.034 | 174.890 | 21% |
| Jumlah | 1,053.656 | 891.876 | 161.779 | 4,022.473 | - |

Keterangan :

- O = Data Asli (data sebenarnya di hari ke-15)
 P = Data hasil prediksi model
 $O - P$ = Kesalahan Prediksi
 n = Jumlah data Pengamatan

Dari Tabel 4.38 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\bullet \quad PE \ 1 = \left| \frac{-13.419}{16.378} \right| \times 100 \%$$

$$PE \ 1 = 81.9 \%$$

$$\bullet \quad PE \ 2 = \left| \frac{-12.214}{21.811} \right| \times 100 \%$$

$$PE \ 2 = 56.0 \%$$

Dilanjutkan sampai PE ke 23

$$\bullet \quad PE \ 23 = \left| \frac{-1.890}{41.117} \right| \times 100 \%$$

$$PE \ 23 = 4,6 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{81.9\% + 56.0\% \dots \dots 4.6\%}{23}$$

$$MAPE = 21 \%$$

- **R²**

Didapatkan langsung dari software R = 0.734

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.734} \\ &= 0.857 \end{aligned}$$

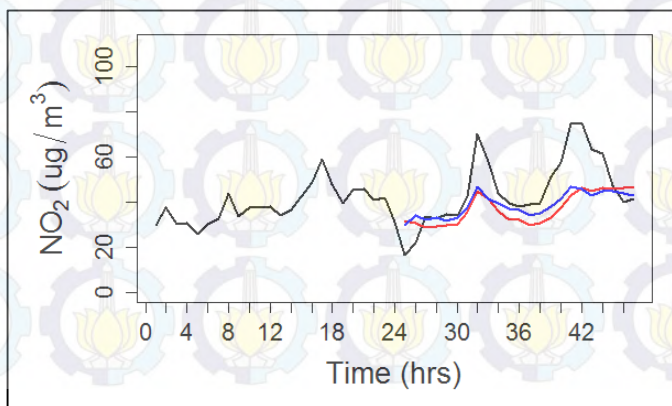
Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.39

Tabel 4.39 Performa Model ARIMA dan STL

| Data In Sample | | 29 Juli 2014 - 12 Agustus 2014 | |
|--------------------------------|----------|--------------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 26% | 21% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.501 | 0.734 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.708 | 0.857 |

Berdasarkan Tabel 4.39 performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL dan MAPE kedua model tidak sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R² dan korelasi. Model ARIMA memiliki nilai R² dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.39 nilai R² STL hanya sebesar 0.501 sedangkan ARIMA 0.734. Besarnya nilai R² dan korelasi pada model ARIMA telah memenuhi standard.

Nilai R² dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.13



Gambar 4. 13 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 1 parameter NO₂

Berdasarkan Gambar 4.13 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada Grafik tersebut dapat dilihat pola Model ARIMA lebih identik dengan data Asli daripada model STL. Hal itu dibuktikan dengan nilai R^2 pada Model STL hanya sebesar 0,501 dan korelasi sebesar 0.708, sedangkan model ARIMA memiliki nilai R^2 0.734 dan nilai korelasi sebesar 0.857.

Jadi dapat disimpulkan pada pengujian pertama, kedua model tidak mampu untuk memprediksi zat pencemar NO₂ meskipun nilai R^2 ARIMA lebih baik, tetapi nilai MAPE masih dibawah standard yang ditetapkan

2. Data Tanggal 1 Agustus 2014 - 15 Agustus 2014

Pengujian model menggunakan data in sample pada tanggal 1 Agustus 2014 hingga tanggal 14 Agustus 2014 (14 hari) dan satu hari digunakan sebagai data verifikasi yaitu pada tanggal 15 Agustus 2014. Karakteristik zat pencemar NO₂ pada tanggal 1 Agustus 2014 hingga tanggal 14 Agustus 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (15 Agustus 2014)

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

Tabel 4.40 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|---------|-----------------------|------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 32.907 | 35.433 | - 2.526 | 6.381 | 7.7% |
| 2 | 34.212 | 34.935 | - 0.723 | 0.523 | 2.1% |

| | | | | | |
|------------------|----------------|----------------|---------------|------------------|------------|
| 3 | 31.143 | 32.909 | - 1.766 | 3.120 | 5.7% |
| 4 | 29.306 | 30.970 | - 1.664 | 2.769 | 5.7% |
| 5 | 21.585 | 30.917 | - 9.332 | 87.088 | 43.2% |
| 6 | 23.487 | 32.132 | - 8.645 | 74.740 | 36.8% |
| 7 | 42.149 | 39.040 | 3.109 | 9.664 | 7.4% |
| 8 | 72.527 | 52.262 | 20.265 | 410.659 | 27.9% |
| 9 | 49.816 | 46.889 | 2.927 | 8.567 | 5.9% |
| 10 | 48.911 | 39.458 | 9.452 | 89.348 | 19.3% |
| 11 | 40.298 | 35.339 | 4.958 | 24.583 | 12.3% |
| 12 | 35.085 | 35.244 | - 0.159 | 0.025 | 0.5% |
| 13 | 33.567 | 32.566 | 1.000 | 1.001 | 3.0% |
| 14 | 33.587 | 33.397 | 0.190 | 0.036 | 0.6% |
| 15 | 45.229 | 36.300 | 8.929 | 79.719 | 19.7% |
| 16 | 52.602 | 41.138 | 11.464 | 131.415 | 21.8% |
| 17 | 62.694 | 47.661 | 15.033 | 225.996 | 24.0% |
| 18 | 64.230 | 50.261 | 13.969 | 195.137 | 21.7% |
| 19 | 55.263 | 46.308 | 8.955 | 80.183 | 16.2% |
| 20 | 45.435 | 46.579 | - 1.144 | 1.309 | 2.5% |
| 21 | 44.514 | 44.136 | 0.378 | 0.143 | 0.8% |
| 22 | 42.964 | 43.593 | - 0.629 | 0.396 | 1.5% |
| 23 | 45.768 | 42.657 | 3.110 | 9.673 | 6.8% |
| Rata-Rata | 42.796 | 39.571 | 3.354 | 62.716 | 13% |
| Jumlah | 941.505 | 910.123 | 77.149 | 1,442.474 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.40 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-2.526}{32.907} \right| \times 100 \%$

$$PE\ 1 = 7.7 \%$$

- $PE\ 2 = \left| \frac{-0.723}{34.212} \right| \times 100 \%$

$$PE\ 2 = 2.1 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{3.110}{45.768} \right| \times 100 \%$

$$PE\ 23 = 6.8 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{7.7\% + 2.1\% \dots \dots 6.8\%}{23}$$

$$MAPE = 13 \%$$

- **R²**

Didapatkan langsung dari software R = 0.838

- **Korelasi**

$$Korelasi = \sqrt{R^2}$$

$$= \sqrt{0.838}$$

$$= 0.915$$

a. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.41 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| | 32.907 | 39.447 | - 6.540 | 42.769 | 19.9% |
| 2 | 34.212 | 41.209 | - 6.997 | 48.952 | 20.5% |
| 3 | 31.143 | 39.841 | - 8.699 | 75.672 | 27.9% |
| 4 | 29.306 | 31.161 | - 1.856 | 3.444 | 6.3% |
| 5 | 21.585 | 31.522 | - 9.937 | 98.737 | 46.0% |
| 6 | 23.487 | 32.809 | - 9.323 | 86.913 | 39.7% |
| 7 | 42.149 | 41.590 | 0.558 | 0.312 | 1.3% |
| 8 | 72.527 | 53.576 | 18.951 | 359.130 | 26.1% |
| 9 | 49.816 | 48.411 | 1.405 | 1.975 | 2.8% |
| 10 | 48.911 | 43.759 | 5.152 | 26.539 | 10.5% |
| 11 | 40.298 | 41.142 | - 0.844 | 0.712 | 2.1% |
| 12 | 35.085 | 42.151 | - 7.066 | 49.931 | 20.1% |
| 13 | 33.567 | 42.152 | - 8.586 | 73.714 | 25.6% |

| | | | | | |
|------------------|----------------|----------------|----------------|------------------|------------|
| 14 | 33.587 | 40.801 | - 7.214 | 52.048 | 21.5% |
| 15 | 45.229 | 42.166 | 3.062 | 9.378 | 6.8% |
| 16 | 52.602 | 47.184 | 5.418 | 29.353 | 10.3% |
| 17 | 62.694 | 53.797 | 8.897 | 79.159 | 14.2% |
| 18 | 64.230 | 52.802 | 11.428 | 130.605 | 17.8% |
| 19 | 55.263 | 49.081 | 6.182 | 38.212 | 11.2% |
| 20 | 45.435 | 46.184 | - 0.749 | 0.561 | 1.6% |
| 21 | 44.514 | 43.058 | 1.456 | 2.120 | 3.3% |
| 22 | 42.964 | 43.386 | - 0.422 | 0.178 | 1.0% |
| 23 | 45.768 | 42.485 | 3.282 | 10.774 | 7.2% |
| Rata-Rata | 42.925 | 43.031 | - 0.106 | 53.095 | 15% |
| Jumlah | 987.273 | 989.713 | - 2.440 | 1,221.188 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.41 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n} ; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE 1 = \left| \frac{- 6.540}{32.907} \right| \times 100 \%$$

$$PE 1 = 19.9 \%$$

$$\blacksquare PE 2 = \left| \frac{- 6.997}{34.212} \right| \times 100 \%$$

$$PE 2 = 20.5 \%$$

Dilanjutkan sampai PE ke 23

$$PE_{23} = \left| \frac{3.282}{45.768} \right| \times 100 \%$$

$$PE_{23} = 7.2 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{19.9\% + 20.5\% \dots \dots 7.2\%}{23}$$

$$MAPE = 15 \%$$

- **R²**

Didapatkan langsung dari software $R = 0.875$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.875} \\ &= 0.935 \end{aligned}$$

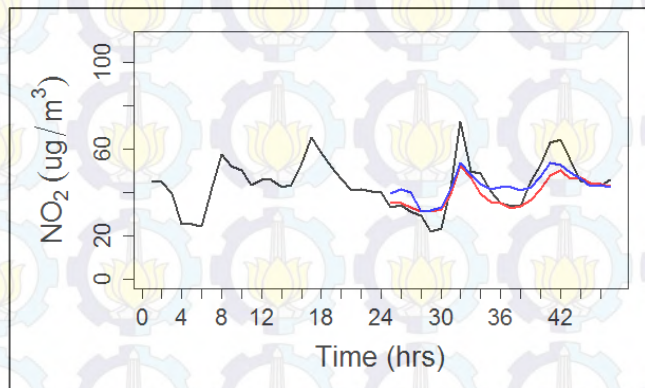
Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.42

Tabel 4.42 Performa Model ARIMA dan STL

| Data In Sample | | 1 Agustus 2014 - 15 Agustus 2014 | |
|--------------------------------|----------|----------------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 13% | 15% |
| R ² | STANDARD | > 0.6 | |
| | HASIL | 0.838 | 0.875 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.915 | 0.935 |

Berdasarkan Tabel 4.42 MAPE kedua model sudah memenuhi standard, dapat dilihat performa STL lebih bagus dibandingkan ARIMA karena MAPE pada Model STL lebih kecil daripada ARIMA dan Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model ARIMA memiliki nilai R^2 dan korelasi yang lebih besar daripada STL, terlihat pada Tabel 4.42 nilai R^2 ARIMasebesar 0.63 sedangkan ARIMA hanya 0.355. Besarnya nilai R^2 dan korelasi pada kedua model tidak memenuhi standard.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.10



Gambar 4. 14 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 2 parameter NO₂

Berdasarkan Gambar 4.14 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada Grafik tersebut dapat dilihat pola Model STL dan ARIMA hamper sama dan mengikuti pola data asli. Hal itu dibuktikan dengan nilai R^2 pada Model STL dan ARIMA yang tidak terpaut jauh, yaitu masing masing sebesar 0.838 dan 0.875.

Jadi dapat disimpulkan pada pengujian kedua, kedua model mampu untuk memprediksi zat pencemar O_3 karena kriteria penilaian pada kedua model sudah memenuhi standard.

3. Data Tanggal 19 Juli 2014 - 31 Juli 2014

Pengujian model menggunakan data in sample pada tanggal 4 Juli 2014 hingga 17 Agustus 2014 (14 hari) dan satu hari digunakan sebagai data verivikasi yaitu pada tanggal 18 Agustus 2014 Karakteristik zat pencemar O_3 pada tanggal 4 Juli 2014 hingga tanggal 17 Agustus 2014 dapat dilihat di lampiran. Setiap Model akan diuji kemampuannya dalam memprediksi hari ke-15 (tanggal 18 Agustus 2014).

Model terbaik adalah model dengan performa sesuai standard kriteria yang telah ditentukan pada bab sebelumnya. Berikut adalah analisis dan hasil prediksi masing masing model

a. Performa Seasonal Trend Decomposition based on Loess (STL)

Data prediksi STL dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Berikut adalah tabel perbandingan data prediksi STL yang dihasilkan dari program R

Tabel 4.43 Performa Model Seasonal Decomposition Based on Loess (STL)

| Seasonal Decomposition Based on Loess (STL) | | | | | |
|---|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 36.361 | 39.409 | - 3.048 | 9.288 | 8.4% |
| 2 | 30.414 | 38.899 | - 8.485 | 72.002 | 27.9% |
| 3 | 29.608 | 37.483 | - 7.875 | 62.023 | 26.6% |
| 4 | 31.259 | 35.155 | - 3.896 | 15.176 | 12.5% |
| 5 | 33.546 | 34.620 | - 1.075 | 1.155 | 3.2% |
| 6 | 36.404 | 36.662 | - 0.258 | 0.067 | 0.7% |
| 7 | 49.640 | 45.608 | 4.032 | 16.254 | 8.1% |
| 8 | 80.132 | 60.337 | 19.795 | 391.855 | 24.7% |
| 9 | 61.491 | 55.162 | 6.329 | 40.054 | 10.3% |

| | | | | | |
|------------------|------------------|------------------|-----------------|------------------|------------|
| 10 | 44.727 | 47.435 | - 2.709 | 7.338 | 6.1% |
| 11 | 41.862 | 42.727 | - 0.866 | 0.750 | 2.1% |
| 12 | 36.075 | 42.472 | - 6.397 | 40.927 | 17.7% |
| 13 | 45.598 | 40.076 | 5.522 | 30.497 | 12.1% |
| 14 | 39.301 | 39.679 | - 0.378 | 0.143 | 1.0% |
| 15 | 44.278 | 43.555 | 0.723 | 0.523 | 1.6% |
| 16 | 44.862 | 49.247 | - 4.385 | 19.232 | 9.8% |
| 17 | 51.629 | 57.337 | - 5.709 | 32.590 | 11.1% |
| 18 | 52.679 | 58.980 | - 6.302 | 39.713 | 12.0% |
| 19 | 52.380 | 55.382 | - 3.003 | 9.018 | 5.7% |
| 20 | 45.583 | 50.544 | - 4.961 | 24.613 | 10.9% |
| 21 | 38.956 | 49.517 | - 10.562 | 111.552 | 27.1% |
| 22 | 40.951 | 49.114 | - 8.163 | 66.641 | 19.9% |
| 23 | 43.664 | 48.155 | - 4.491 | 20.166 | 10.3% |
| Rata-Rata | 43.974 | 45.981 | - 2.007 | 43.982 | 12% |
| Jumlah | 1,011.394 | 1,057.556 | - 46.162 | 1,011.575 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.43 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

$$\blacksquare PE_1 = \left| \frac{-3.048}{36.361} \right| \times 100 \%$$

$$PE_1 = 8.4 \%$$

- $PE\ 2 = \left| \frac{-8.485}{30.414} \right| \times 100\ %$

$$PE\ 2 = 27.9\ %$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-4.491}{43.664} \right| \times 100\ %$

$$PE\ 23 = 10.3\ %$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{8.4\% + 27.9\% \dots \dots 10.3\%}{23}$$

$$MAPE = 12\ %$$

- R^2

Didapatkan langsung dari software $R = 0.683$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.683} \\ &= 0.826 \end{aligned}$$

b. Performa Autoregressive Integrated Moving Average (ARIMA)

Data prediksi ARIMA dibandingkan dengan data Verifikasi untuk melihat nilai MAPE, R^2 dan Korelasi. Perhitungan ARIMA berbeda dengan STL dikarenakan model ARIMA harus dilihat stasioneritasnya terlebih dahulu. Identifikasi stasioneritas terhadap mean dapat dilakukan secara visual dengan menggunakan time series plot dan plot ACF. Sedangkan identifikasi stasioneritas terhadap varians dapat dilakukan dengan plot Box-Cox. Time

series plot dari data pencemar tsb. Berdasarkan Plot ACF dan PACF pada Lampiran A. data dinyatakan sudah stasioner terhadap mean. Setelah itu dilakukan Transformasi Box-Cox untuk melihat data telah stasioner terhadap varians. Hasilnya adalah besar lambda pada pengujian ini adalah 1 lambda tersebut dimasukkan kedalam perintah auto arima, perintah tersebut langsung menentukan orde terbaik ARIMA (musiman atau Non Musiman)

Berikut adalah tabel perbandingan data prediksi ARIMA yang dihasilkan dari program R :

Tabel 4.44 Performa Model Autoregressive Integrated Moving Average (ARIMA)

| Autoregressive Integrated Moving Average (ARIMA) | | | | | |
|--|--------|--------|---------|-----------------------|-------|
| No | O | P | O - P | (O - P) ² | PE |
| 1 | 36.361 | 37.495 | - 1.134 | 1.286 | 3.1% |
| 2 | 30.414 | 34.659 | - 4.245 | 18.021 | 14.0% |
| 3 | 29.608 | 33.920 | - 4.312 | 18.595 | 14.6% |
| 4 | 31.259 | 32.543 | - 1.284 | 1.649 | 4.1% |
| 5 | 33.546 | 31.120 | 2.426 | 5.883 | 7.2% |
| 6 | 36.404 | 32.092 | 4.311 | 18.587 | 11.8% |
| 7 | 49.640 | 39.446 | 10.194 | 103.924 | 20.5% |
| 8 | 80.132 | 51.440 | 28.692 | 823.208 | 35.8% |
| 9 | 61.491 | 55.717 | 5.773 | 33.329 | 9.4% |
| 10 | 44.727 | 45.850 | - 1.123 | 1.262 | 2.5% |
| 11 | 41.862 | 39.704 | 2.157 | 4.654 | 5.2% |
| 12 | 36.075 | 38.880 | - 2.805 | 7.866 | 7.8% |
| 13 | 45.598 | 37.428 | 8.170 | 66.753 | 17.9% |
| 14 | 39.301 | 37.314 | 1.987 | 3.948 | 5.1% |
| 15 | 44.278 | 39.715 | 4.563 | 20.825 | 10.3% |
| 16 | 44.862 | 46.727 | - 1.865 | 3.478 | 4.2% |
| 17 | 51.629 | 55.680 | - 4.052 | 16.418 | 7.8% |
| 18 | 52.679 | 57.533 | - 4.854 | 23.565 | 9.2% |
| 19 | 52.380 | 55.059 | - 2.679 | 7.178 | 5.1% |
| 20 | 45.583 | 40.983 | 4.599 | 21.154 | 10.1% |
| 21 | 38.956 | 41.404 | - 2.449 | 5.997 | 6.3% |
| 22 | 40.951 | 43.017 | - 2.067 | 4.272 | 5.0% |

| | | | | | |
|------------------|------------------|----------------|---------------|------------------|------------|
| 23 | 43.664 | 45.223 | - 1.559 | 2.431 | 3.6% |
| Rata-Rata | 43.974 | 42.302 | 1.672 | 52.795 | 10% |
| Jumlah | 1,011.394 | 972.949 | 38.445 | 1,214.282 | - |

Keterangan :

O = Data Asli (data sebenarnya di hari ke-15)

P = Data hasil prediksi model

$O - P$ = Kesalahan Prediksi

n = Jumlah data Pengamatan

Dari Tabel 4.44 dapat diketahui

- **MAPE (Mean Absolute Percentage Error)**

Contoh perhitungan pada No. 1

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}; PE = \left| \frac{O-P}{O} \right| \times 100 \%$$

- $PE\ 1 = \left| \frac{-1.134}{36.361} \right| \times 100 \%$

$$PE\ 1 = 3.1 \%$$

- $PE\ 2 = \left| \frac{-4.245}{30.414} \right| \times 100 \%$

$$PE\ 2 = 14.0 \%$$

Dilanjutkan sampai PE ke 23

- $PE\ 23 = \left| \frac{-1.559}{43.664} \right| \times 100 \%$

$$PE\ 23 = 3.6 \%$$

$$MAPE = \frac{\sum_{i=1}^n PE_i}{n}$$

$$MAPE = \frac{3.1\% + 14.0\% \dots \dots 3.6\%}{23}$$

$$MAPE = 10 \%$$

- R^2

Didapatkan langsung dari software $R = 0.586$

- **Korelasi**

$$\begin{aligned} \text{Korelasi} &= \sqrt{R^2} \\ &= \sqrt{0.586} \\ &= 0.776 \end{aligned}$$

Performa kedua model pada pengujian pertama dapat dilihat melalui Tabel 4.45

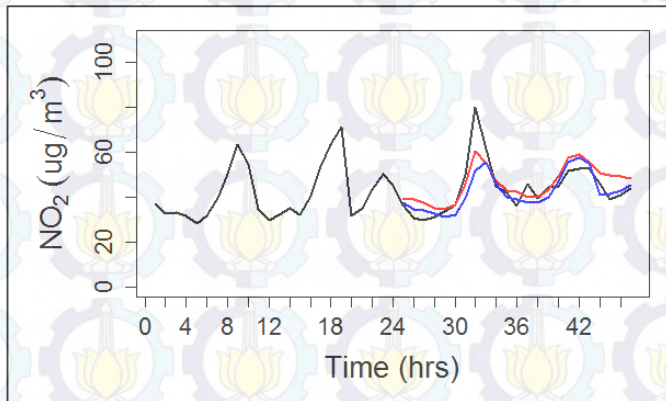
Tabel 4.45 Performa Model ARIMA dan STL

| Data In Sample | | 4 Juli 2014 - 18 Agustus 2014 | |
|--------------------------------|----------|-------------------------------|-------|
| Model | | STL | ARIMA |
| Mean Absolute Percentage Error | STANDARD | <20% | |
| | HASIL | 12% | 10% |
| R^2 | STANDARD | > 0.6 | |
| | HASIL | 0.683 | 0.586 |
| Correlation | STANDARD | (+ 0.774) | |
| | HASIL | 0.826 | 0.766 |

Berdasarkan Tabel 4.45 performa ARIMA lebih bagus dibandingkan STL karena MAPE pada Model ARIMA lebih kecil daripada STL dan MAPE kedua model tidak sesuai dengan standard yang telah ditetapkan. Penilaian performa model juga dibuktikan dengan nilai R^2 dan korelasi. Model ARIMA memiliki nilai R^2 dan korelasi yang lebih kecil daripada STL, terlihat pada Tabel 4.39 nilai R^2 STL sebesar 0.683 sedangkan ARIMA 0.586. Oleh

sebab itu model STL memiliki nilai R^2 dan korelasi pada model yang memenuhi standard.

Nilai R^2 dan korelasi menunjukkan kesesuaian prediksi model terhadap data Verifikasi (data asli). Hal ini dibuktikan melalui Gambar grafik 4.13



Gambar 4. 15 Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 3 parameter NO₂

Berdasarkan Gambar 4.15 dapat dilihat kesesuaian pola grafik data hasil prediksi model STL dan ARIMA terhadap data verifikasi. Grafik Warna Merah menunjukkan grafik hasil prediksi model STL dan grafik yang berwarna biru merupakan grafik hasil permodelan ARIMA, sedangkan grafik yang berwarna hitam adalah data verifikasi (data asli). Pada Grafik tersebut dapat dilihat pola Model STL dan ARIMA hampir sama dan mengikuti pola data asli. Hal itu dibuktikan dengan nilai R^2 pada Model STL dan ARIMA yang tidak terpaut jauh, yaitu masing masing sebesar 0.683 dan 0.586

Jadi dapat disimpulkan pada ketiga pengujian kedua model tidak mampu memprediksi zat pencemar NO₂, karena kedua model tidak menunjukkan konsistensi dalam memprediksi zat pencemar tersebut.

4.2. Performa Model STL dan Model ARIMA terhadap data yang hilang

Performa Model Seasonal Decomposition based on Loess (STL) dan ARIMA Juga dapat ditentukan dengan kemampuan tiap model dalam mencari data yang hilang pada waktu yang sama. Jumlah data yang hilang di variasikan menjadi dua macam yaitu 15 % dan 25 %, yang masing masing memiliki tipe hilang yang berbeda yakni beruntun dan acak.

4.2.1. SO₂

Performa model STL dan ARIMA terhadap data yang hilang dibagi menjadi dua skenario, yaitu data hilang secara beruntun dan data hilang secara acak. Berikut adalah hasil dan analisis data dari kedua model.

a. Data Hilang Berurutan

Untuk

Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4.46 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun

| Data In Sample | | 10 Agustus 2014 - 24 Agustus 2014 | | | |
|-------------------------------|----------|-----------------------------------|-------|--------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20 % | | | |
| | HASIL | 39% | 32% | 102% | 19% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 6.912 | 6.519 | 11.727 | 7.990 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.6912 | 0.737 | 0.626 | 0.683 |
| Correlation | STANDARD | 0.774 | | | |
| | HASIL | 0.831 | 0.859 | 0.791 | 0.826 |

Berdasarkan tabel 4.46 Model R^2 pada tiap Model cenderung meningkat dan Mean Absolute Percentage Error pada model ARIMA menurun drastis. Hal ini menandakan model ARIMA memiliki kesalahan prediksi yang sangat besar dibandingkan dengan STL yang perbedaan kesalahan pada prediksi dari 2 tipe data hilang tidak terlalu besar (39% dan 32%)

b. Data Hilang Acak

Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 47 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak

| Data In Sample | | 10 Agustus 2014 - 24 Agustus 2014 | | | |
|-------------------------------|----------|-----------------------------------|-------|-------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20 % | | | |
| | HASIL | 29% | 33% | 72% | 27% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 6.557 | 7.602 | 9.369 | 7.668 |
| R^2 | STANDARD | > 0.6 | | | |
| | HASIL | 0.703 | 0.629 | 0.682 | 0.589 |
| Correlation | STANDARD | 0.774 | | | |
| | HASIL | 0.838 | 0.793 | 0.826 | 0.767 |

Berdasarkan tabel 4.47, Model STL dengan data hilang sebesar 15 % memiliki nilai Mean Absolut Percentage Error (MAPE) yang lebih kecil dibandingkan model ARIMA yang memiliki nilai MAPE sebesar 72 %. Dan perubahan nilai MAPE pada model STL masih lebih kecil dibandingkan perubahan nilai MAPE pada ARIMA. Hal ini membuktikan bahwa tingkat kesalahan prediksi ARIMA lebih besar daripada STL. Jadi dapat disimpulkan bahwa STL lebih baik daripada ARIMA pada data yang dihilangkan secara

acak. Untuk menguatkan pernyataan ini perlu adanya pengujian untuk parameter lain.

4.2.2. CO

Performa model STL dan ARIMA terhadap data yang hilang dibagi menjadi dua skenario, yaitu data hilang secara beruntun dan data hilang secara acak. Berikut adalah hasil dan analisis data dari kedua model.

a. Data Hilang Berurutan

Pengujian terhadap data hilang untuk zat pencemar CO dilakukan untuk melihat konsistensi data hilang di berbagai zat pencemar. Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 48 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun

| Data In Sample | | 15 Juli 2014 - 29 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|-------|-------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20 % | | | |
| | HASIL | 44% | 34% | 8% | 15% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 0.125 | 0.097 | 0.033 | 0.053 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.82 | 0.898 | 0.898 | 0.648 |
| Correlation | STANDARD | 0.774 | | | |
| | HASIL | 0.906 | 0.948 | 0.948 | 0.805 |

Berdasarkan tabel 4.48 Model R² pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL cukup besar. Hal ini disebabkan Karena data hasil prediksi dari Model STL bernilai negatif (-) sedangkan data verifikasinya tidak demikian (hasil prediksi dan verifikasi dapat dilihat pada lampiran)

Jadi pada pengujian data hilang secara beruntun untuk parameter CO, Model STL dinilai tidak mampu memprediksi dengan baik.

b. Data Hilang Acak

Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 49 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak

| Data In Sample | | 15 Juli 2014 - 29 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|-------|-------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20 % | | | |
| | HASIL | 75% | 61% | 7% | 16% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 0.203 | 0.167 | 0.028 | 0.073 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.824 | 0.837 | 0.914 | 0.890 |
| Correlation | STANDARD | 0.774 | | | |
| | HASIL | 0.908 | 0.915 | 0.956 | 0.943 |

Berdasarkan tabel 4.49 Model R² pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL cukup besar. Hal ini disebabkan Karena data hasil prediksi dari Model STL bernilai negatif (-) sedangkan data verifikasinya tidak demikian (hasil prediksi dan verifikasi dapat dilihat pada lampiran). Jadi pada pengujian data hilang secara acak untuk parameter CO, Model STL dinilai tidak mampu memprediksi dengan baik.

4.2.3. PM₁₀

Performa model STL dan ARIMA terhadap data yang hilang dibagi menjadi dua skenario, yaitu data hilang secara beruntun dan data hilang secara acak. Berikut adalah hasil dan analisis data dari kedua model.

a. Data Hilang Berurutan

Pengujian terhadap data hilang untuk zat pencemar PM₁₀ dilakukan untuk melihat konsistensi data hilang di berbagai zat pencemar. Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 50 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|--------|--------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20% | | | |
| | HASIL | 69% | 57% | 23% | 26% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 26.056 | 20.217 | 10.859 | 7.992 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.522 | 0.467 | 0.788 | 0.422 |
| Correlation | STANDARD | (+ 0.774) | | | |
| | HASIL | 0.722 | 0.683 | 0.888 | 0.65 |

Berdasarkan tabel 4.50 Model R² pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL cukup besar. Hal ini disebabkan Karena pengaruh interpolasi formula untuk mengisi data yang hilang. (hasil prediksi dan verifikasi dapat dilihat pada lampiran) Jadi pada pengujian data hilang secara beruntun untuk parameter PM₁₀, Model STL dan ARIMA dinilai tidak mampu memprediksi dengan baik.

b. Data Hilang Acak

Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 51 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|--------|--------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20% | | | |
| | HASIL | 63% | 60% | 76% | 27% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 19.914 | 19.476 | 23.158 | 9.454 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.423 | 0.437 | 0.65 | 0.113 |
| Correlation | STANDARD | 0.5 | | | |
| | HASIL | 0.65 | 0.661 | 0.806 | 0.336 |

Berdasarkan tabel 4.51 Model R² pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL cukup besar. Hal ini disebabkan Karena pengaruh interpolasi formula untuk mengisi data yang hilang. (hasil prediksi dan verifikasi dapat dilihat pada lampiran) Jadi pada pengujian data hilang secara acak untuk parameter PM₁₀, Model STL dan ARIMA dinilai tidak mampu memprediksi dengan baik.

4.2.4. O₃

Performa model STL dan ARIMA terhadap data yang hilang dibagi menjadi dua skenario, yaitu data hilang secara beruntun dan data hilang secara acak. Berikut adalah hasil dan analisis data dari kedua model.

a. Data Hilang Berurutan

Pengujian terhadap data hilang untuk zat pencemar O₃ dilakukan untuk melihat konsistensi data hilang di berbagai zat pencemar. Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 52 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|--------|--------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20% | | | |
| | HASIL | 69% | 57% | 23% | 26% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 26.056 | 20.217 | 10.859 | 7.992 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.522 | 0.467 | 0.788 | 0.422 |
| Correlation | STANDARD | (+ 0.774) | | | |
| | HASIL | 0.722 | 0.683 | 0.888 | 0.65 |

Berdasarkan tabel 4.52 Model R² pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL cukup besar. Hal ini disebabkan Karena pengaruh interpolasi formula untuk mengisi data yang hilang. (hasil prediksi dan verifikasi dapat dilihat pada lampiran) Jadi pada pengujian data

hilang secara beruntun untuk parameter O_3 , Model STL dan ARIMA dinilai tidak mampu memprediksi dengan baik.

b. Data Hilang Acak

Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 53 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|-------|-------|--------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20% | | | |
| | HASIL | 15% | 15% | 19% | 21% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 6.448 | 6.61 | 9.328 | 10.678 |
| R^2 | STANDARD | > 0.6 | | | |
| | HASIL | 0.905 | 0.909 | 0.65 | 0.818 |
| Correlation | STANDARD | (+ 0.774) | | | |
| | HASIL | 0.951 | 0.953 | 0.806 | 0.904 |

Berdasarkan tabel 4.53 Model R^2 pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL cukup besar. Hal ini disebabkan Karena pengaruh interpolasi formula untuk mengisi data yang hilang. (hasil prediksi dan verifikasi dapat dilihat pada lampiran) Jadi pada pengujian data hilang secara acak untuk parameter O_3 , Model STL dan ARIMA dinilai tidak mampu memprediksi dengan baik.

4.2.5. NO₂

Performa model STL dan ARIMA terhadap data yang hilang dibagi menjadi dua skenario, yaitu data hilang secara beruntun dan data hilang secara acak. Berikut adalah hasil dan analisis data dari kedua model.

a. Data Hilang Berurutan

Pengujian terhadap data hilang untuk zat pencemar NO₂ dilakukan untuk melihat konsistensi data hilang di berbagai zat pencemar. Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 54 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Beruntun

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|-------|-------|-------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20% | | | |
| | HASIL | 14% | 17% | 16% | 14% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 9.173 | 9.992 | 7.418 | 7.197 |
| R ² | STANDARD | > 0.6 | | | |
| | HASIL | 0.805 | 0.843 | 0.745 | 0.854 |
| Correlation | STANDARD | (+ 0.774) | | | |
| | HASIL | 0.897 | 0.918 | 0.863 | 0.924 |

Berdasarkan tabel 4.54 Model R² pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada model STL sudah memenuhi standard. Hal ini disebabkan karena pengaruh interpolasi pada data (hasil prediksi dan verifikasi dapat dilihat pada lampiran) Jadi pada pengujian data hilang secara beruntun

untuk parameter NO_2 , kedua Model dinilai tidak mampu memprediksi dengan baik.

b. Data Hilang Acak

Berikut adalah performa STL dan ARIMA terhadap Data hilang secara beruntun.

Tabel 4. 55 Performa Model STL dan ARIMA Terhadap Data yang Hilang secara Acak

| Data In Sample | | 12 Juli 2014 - 26 Juli 2014 | | | |
|-------------------------------|----------|-----------------------------|--------|-------|--------|
| Model | | STL | | ARIMA | |
| Data Hilang | | 15% | 25% | 15% | 25% |
| Mean Absolut Percentage Error | STANDARD | <20% | | | |
| | HASIL | 13% | 36% | 15% | 25% |
| Root Mean Square Error | STANDARD | - | | | |
| | HASIL | 8.19 | 17.023 | 6.892 | 12.811 |
| R^2 | STANDARD | > 0.6 | | | |
| | HASIL | 0.815 | 0.826 | 0.906 | 0.892 |
| Correlation | STANDARD | (+ 0.774) | | | |
| | HASIL | 0.903 | 0.909 | 0.952 | 0.944 |

Berdasarkan tabel 4.55 Model R^2 pada tiap Model tidak stabil dan Mean Absolute Percentage Error pada kedua model juga tidak stabil. Hal ini disebabkan karena pengaruh interpolasi pada data sehingga data menjadi tidak stabil (hasil prediksi dan verifikasi dapat dilihat pada lampiran). Jadi pada pengujian data hilang secara Acak untuk parameter NO_2 , Model ARIMA lebih mampu memprediksi dengan baik daripada model STL

BAB 5

KESIMPULAN DAN SARAN

- *Kesimpulan*

1. *Model STL dan ARIMA merupakan model terbaik yang digunakan untuk memprediksi CO dan SO₂*
2. *Hilangnya data berpengaruh terhadap hasil prediksi karena pengaruh interpolasi pada data hilang tersebut.*

- *Saran*

1. *Data Training yang digunakan sebaiknya lebih dari 14 hari dan digunakan untuk memprediksi dalam jangka waktu yang lebih lama*
2. *Model yang digunakan harus seasonal agar mempermudah dalam memprediksi dan data hasil prediksi lebih valid*
3. *Pencarian data yang hilang akan lebih baik jika dilakukan dengan metode peramalan bukan dengan metode Interpolasi*

DAFTAR PUSTAKA

Achmad, Rukaesih. 2004. *Kimia Lingkungan*. Yogyakarta: Penerbit ANDI

Badan Lingkungan Hidup. 2009. **Laporan Status Lingkungan Hidup Daerah**. Surabaya.

Cleveland, R.B., Cleveland, W.S., MCRAE, J.E., & TERPENNING, I. (1990)

STL: a seasonal-trend decomposition procedure based on Loess (with discussion), *Journal of Official Statistics*, 6, 3–73.
Hendranata, Anton. 2003. **ARIMA (Autoregressive Moving Average)**, Manajemen Keuangan Sektor Publik FEUI

Lee, C. and Ko, C. .2011. **Short-term Load Forecasting Using Lifting Scheme and ARIMA Models**. *Expert Systems with Applications*, Vol. 38, pp. 5902-5911.

Ihaka, R. dan Gentleman, R. 1996. **R: A Language for Data Analysis and Graphics**. *Journal of Computational and Graphical Statistics* 5 (3): 299–314.

Makridakis, Spyros., Wheelwright, Steven. C., McGee, Victor. E. 1999. **Metode dan Aplikasi Peramalan Jilid 1**. Erlangga. Jakarta.
Makridakis, Spyros., Wheelwright, Steven. C., McGee, Victor. E. 1999. **Metode dan Aplikasi Peramalan Jilid 2**. Interaksana. Batam.

Selvia, dkk., 2011 *Hubungan kadar HbCO dengan Kapasitas Vital Paru-paru pedagang di terminal bus Purwokerto*

Suhartono. 2008. **Analisis Data Statistik dengan R**. Jurusan Statistik, FMIPA-ITS. Surabaya.

Putut, E.L and Widodo, B. 2011 **SIMULASI MODEL DISPERSI POLUTAN KARBON MONOKSIDA DI PINTU MASUK TOL (Studi Kasus Line Source Di Ruas Tol Dupak, Surabaya)**

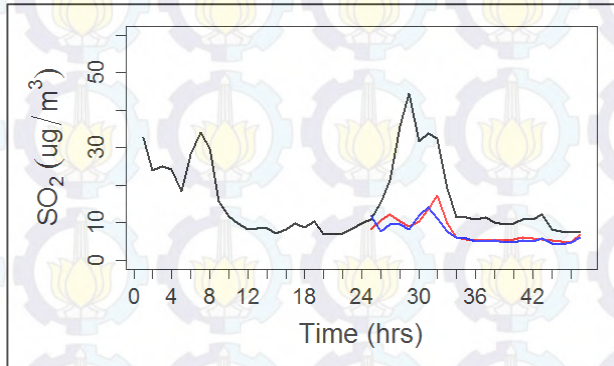
Weir, William. W. S. 2006. **Time Series Analysis : Univariate and Multivariate Methods Second Edition.** Pearson Education, Inc. United States of America

Yulfida Y., Marsaulina I., Ashar T. 2012 (**Perbandingan Kadar Karbon Monoksida (Co) Dan Nitrogen Dioksida (No₂) Di Udara Ambien Berdasarkan Keberadaan Pohon Angsana (Pterocarpus Indicus) Di Beberapa Jalan Raya Di Kota Medan Tahun 2012).** Fakultas Kesehatan Masyarakat. Universitas Sumatera Utara

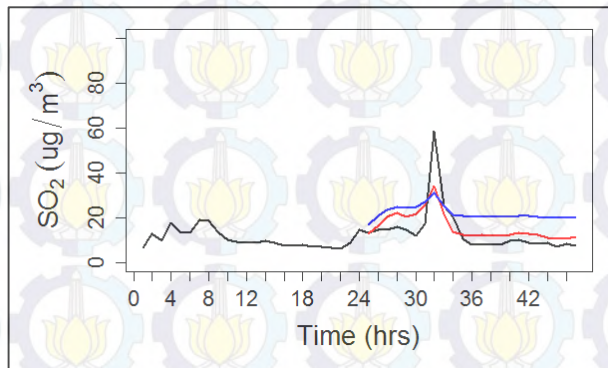


LAMPIRAN A

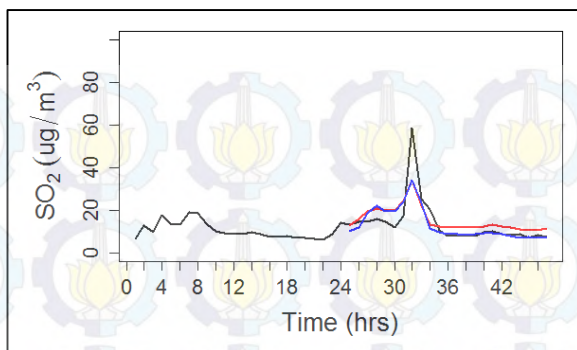
Gambar Grafik Pengaruh Data Hilang



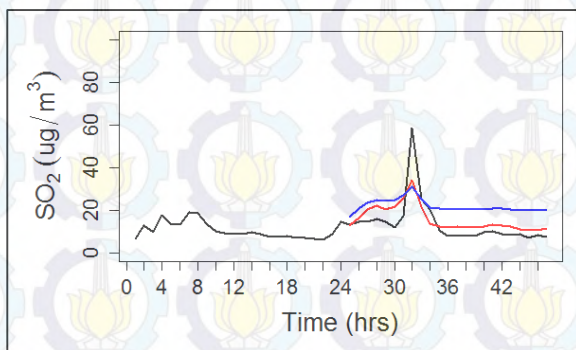
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang acak parameter SO_2



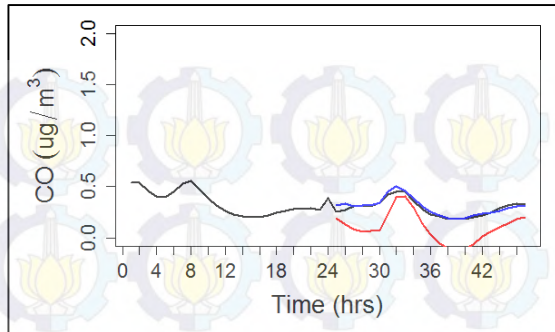
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang acak parameter SO_2



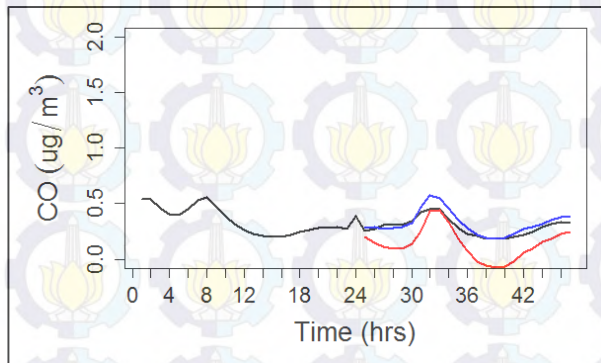
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Runtun parameter SO_2



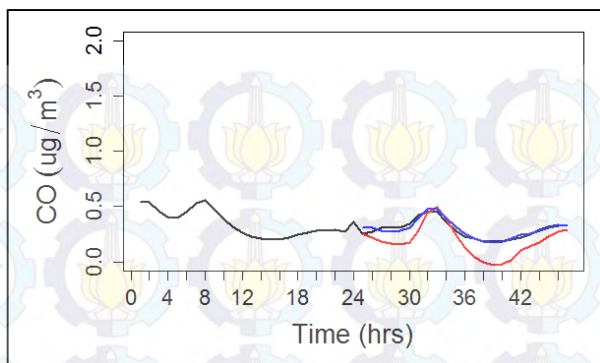
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Runtun parameter SO_2



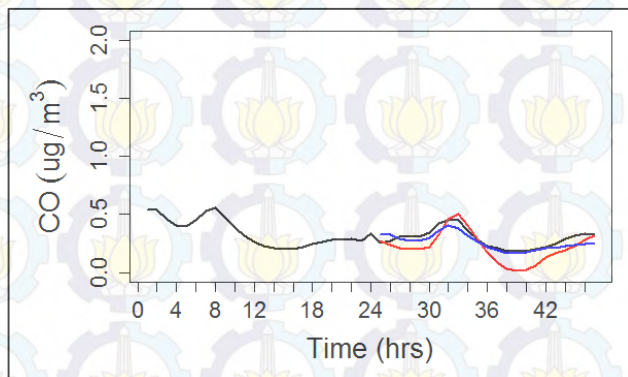
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang acak parameter CO



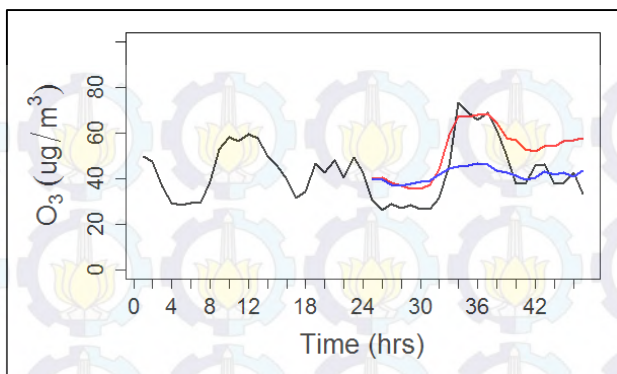
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang acak parameter CO



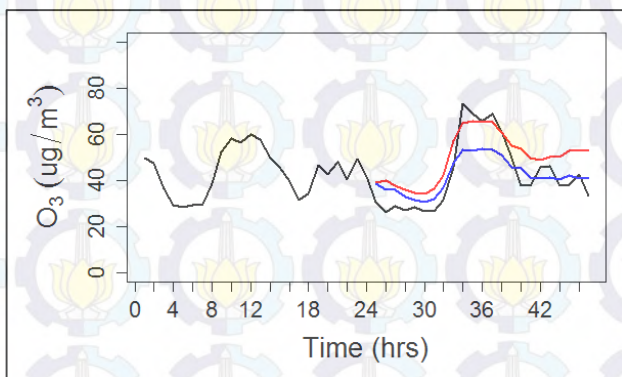
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Runtun parameter CO



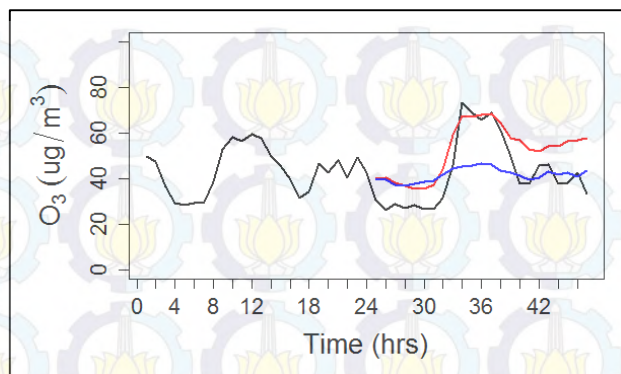
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang runtun parameter CO



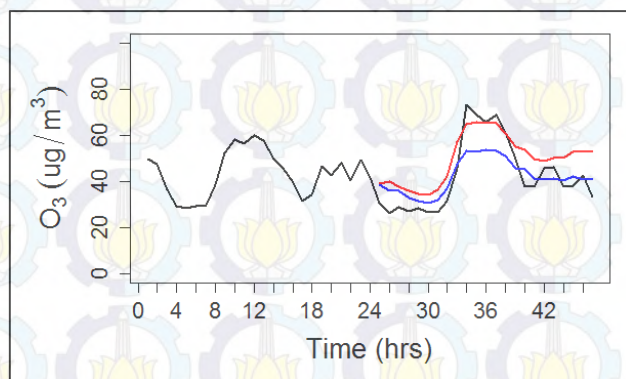
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Runtun parameter O_3



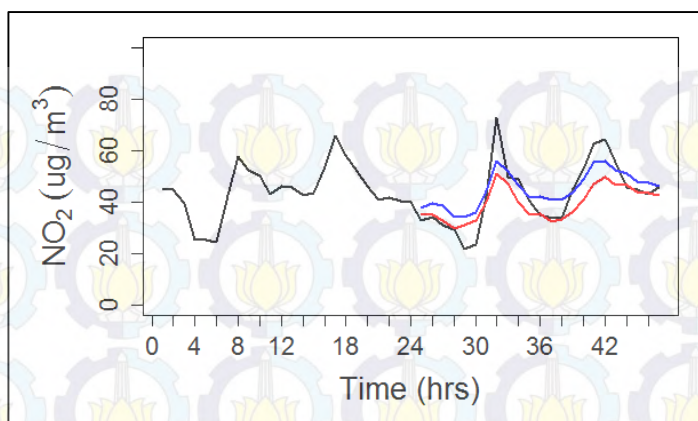
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang Runtun parameter O_3



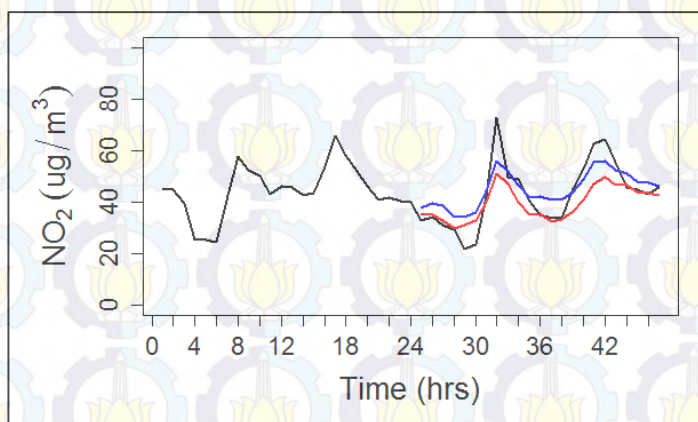
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Acak parameter O_3



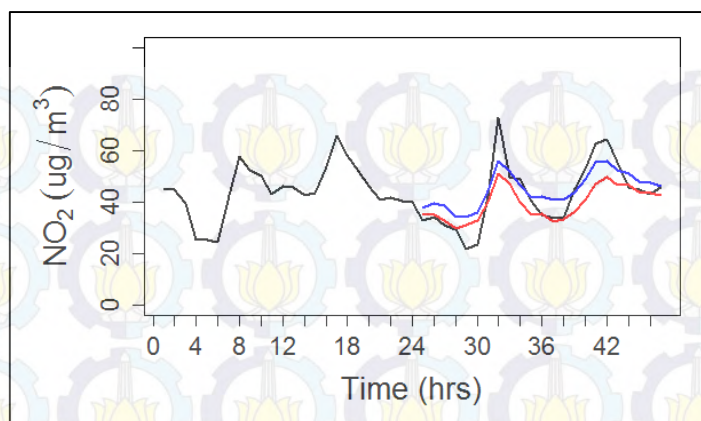
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang Acak parameter O_3



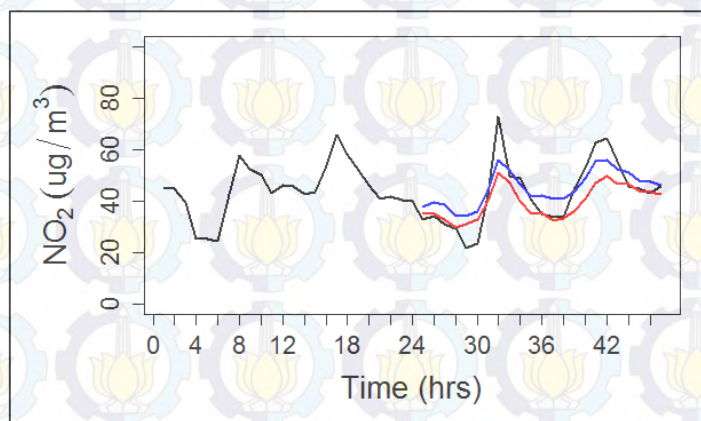
Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Runtun parameter NO_2



Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang Runtun parameter NO_2



Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 15% data hilang Acak parameter NO_2



Grafik permodelan STL dan ARIMA terhadap data Verifikasi pada pengujian 25% data hilang Acak parameter NO_2

BIOGRAFI PENULIS



Penulis dilahirkan di Jakarta pada tanggal 26 september 1993. Pria yang sering dipanggil Ares ini telah menempuh pendidikan formal pada tahun 1998 – 2000 di TK Teratai Bakti . Kemudian dilanjutkan di SDN Purwantoro 03 Pagi Jakarta Timur pada tahun 1999 – 2005 dan dilanjutkan di SMPN 222 Jakarta Timur pada tahun 2005 – 2008. Pendidikan tingkat atas penulis dilalui di SMAN 48 Jakarta Timur pada tahun 2008 – 2011. Pada tahun 2011, penulis melanjutkan kuliah S-1 di

Teknik Lingkungan FTSP ITS dan terdaftar dengan NRP 3311 100 113. Penulis pernah aktif sebagai pengurus Himpunan Mahasiswa Teknik Lingkungan (HMTL) FTSP ITS pada periode tahun 2012 – 2013 dan 2013 – 2014. Penulis pernah menjadi mengikuti beberapa seminar dan pelatihan yang diselenggarakan di tingkat Jurusan ataupun Institut. Pada tahun 2014, penulis melaksanakan Kerja Praktik di Pertamina MOR IV Semarang, Jawa Tengah. Penulis dapat dihubungi via email ahmad.haritsah@yahoo.com.